数学概貌丛书

微积分大意

徐利治著

科学技术文献出版社

数学概貌丛书

微积分大意

徐利治 著

科学技术文献出版社

数学概貌丛书

微积分大意

徐利治 著

*

科学技术文献出版社出版

（北京市复兴路15号）

上海市新华印刷厂印刷

新华书店上海发行所发行　各地新华书店经售

*

开本 850×1156　1/32　印张 8　字数 52,000

1989 年 4 月第 1 版　1989 年 4 月第 1 次印刷

印数 1—4,000 本

定价：1.35 元

ISBN 7-5023-0651-X/O·41

目　　录

引　言

我们假定读者已经具备初等微积分基础知识，在此基础上，本书将着重阐明微积分中的主要问题、基本思想、重要结论、典型方法和有关历史.限于篇幅，本书只能有选择地对某些最基本的事项(包括概念、定理、证明、技巧等)给予详细介绍．至于其余部分，则希望读者能从本书的概略叙述中获得启发而有所领悟.

下面先谈一点微积分的发展简史，以有助于了解它的来龙去脉.

微积分又称“数学分析”，人们还常简单地称之为“分析学”．事实上，“数学分析”是在微积分发展趋于成熟时期才比较通用的名称．它主要包括实数理论、极限理论、微分学、积分学和无穷级数等部分.微分学与积分学是以极限论作为基础的，而极限论又以实数理论为基础．所以，按照顺序，理应先谈实数理论，后讲极限理论，再在此基础上展开微积分学的讨论．然而，正如大家所知道的，牛顿(I. Newton)和莱布尼茨(G. W. Leibniz)大约是在17世纪70年代前后就发现了微积分学基本定理，由此开创了一门新的数学；可是，实数系的逻辑基础却迟至19世纪后叶才建立起来，竟比微积分的发明晚了整整两个世纪.这可以说是数学史上最令人惊奇的事实之一.

大家知道，在数学的学习过程中，凭借直观去学习运算是容易的；但要学习分析和论证，明白为什么可以那样去运算，就比较困难了．微积分学的成长发展过程也是这样．最初，人们（包括牛顿和莱布尼茨在内）大多是借助于直观和经验去引进一系列计算法则，并用以确实解决了一系列实际应用问题，由此更增强了对计算法则正确性的信念．至于有关计算法则赖以成立的原理及其论证，那就完成得很晚了．

举例来说，根据算术运算经验和几何直观，一般人都不会怀疑如下的运算结果：

$$\sqrt{2}\cdot\sqrt{3}=\sqrt{3}\cdot\sqrt{2}=\sqrt{6}.$$

但正如戴德金(J. W. R. Dedekind) 于 1872 年发表的著作中所说的："回忆 1858 年讲授微积分课程时碰到如上的实数运算，至少直到那时候还没有人将它严格证明过．"因为这里出现了无理数运算为何仍然遵循算术运算规律的问题．

谈到导数（微商）的计算法则，涉及到微分学的基础．对此，首先发难的是 18 世纪著名的英国主教贝克莱 (G. Berkeley)．他在 1734 年出版的《分析学家》一书中，对牛顿的'流数术'（即求导数方法）进行了一系列责难．其核心问题，就是要否定导数算法的逻辑合理性．举例来说，为求函数 $x=t^2$ 的导数．按流数术，先给 t 添加一个'流量'（无穷小增量）0，相应地，x 便从 t^2 变为$(t+0)^2$，故 x 获得的流量为 $(t+0)^2-t^2=2t\cdot 0+0^2$．计算两者的比值：

$$\frac{(t+0)^2-t^2}{0}=\frac{2t\cdot 0+0^2}{0}=2t+0.$$

最后，再在所得结果中把 0 项舍掉. 这样便得到牛顿所说的'终极比' (Ultimate ratio)：

$$\dot{x}=2t.$$

显然，这就是 x 关于 t 的导数. 如采用莱布尼茨的记法，也可写成 $\frac{dx}{dt}=2t$.

贝克莱的责难是这样的：首先要问上述算法中的增量 0 究竟是非零量还是真正零？如是非零量，则 $2t+0$ 式中的 0 就不能舍掉；如果是真正零，则 $(t+0)^2$ 与 t^2 无异，即 $(t+0)^2\equiv t^2$，因而所求比值变为无意义的 $\frac{0}{0}$. "总之，不论怎样看，牛顿的流数算法是不合逻辑的." 这就是历史上著名的"贝克莱悖论". 当年，贝克莱还把牛顿的流数(即导数)讥之为"消逝量的鬼魂".

采用如今一般教科书中的记法，牛顿的算法相当于取极限

$$\lim_{\Delta t\to 0}\frac{(t+\Delta t)^2-t^2}{\Delta t}=\lim_{\Delta t\to 0}(2t+\Delta t)=2t.$$

这里应该指出，牛顿本人的概念思维其实是十分清晰的：他以物质运动（包括时间流逝）的连续变化量为直观背景，把增量 0 看成是一个流动量(即变数)，它可以从非零变化到真正零，从而得到他所要求的"终极比". 所谓终极比，即包含有比值(变量)跟着变化而最终达到极限的意思（注意，由于牛顿的流数概念常以力学为背景，故变量

的变化过程总能达到极限值）．所以，牛顿的头脑里不仅有了变量概念，而且还有让变量变化到终极状态时的“极限”概念．

由上所论，可见“贝克莱悖论”之所以出现，并不是由于什么推理上的粗心大意，而恰恰是因为贝克莱之流根本上缺乏变量极限观念，以致只能从初等数学的形式演算上去考虑问题的缘故．

整个18世纪，由于欧拉(L. Euler)和拉格朗日(J. L. Lagrange）等人作了大量工作，微积分在各个应用领域(尤其是在力学科学中)都取得了辉煌成果．虽然出现了贝克莱之流对微分学基础的攻击，却丝毫也没有能动摇人们对微积分功效的信念．正如当年达朗贝尔(J. L. R. D'Alembert)所说的："向前进，你就会产生信心！"

但是，奠基在“无穷小分析法”上的这门新兴学科，毕竟在基本概念方面存在着许多模糊不清之处．事实上，“无穷小”(又称“无限小”，或“无穷小量”）在实数系中是不存在的，把它当作一般实数来运算当然是缺乏根据的；而在当时的一些微积分教材中，却到处借助于无穷小去推导或猜想出一些定理和公式．尽管推导出的结论常常是对的，但缺乏合乎逻辑的证明．

为了避免使用无穷小推理和当时还不甚明确的极限概念，拉格朗日曾试图把整个微积分建立在“泰勒(B. Taylor)展开式”的基础上．他先用代数方法证明泰勒展开式，接着把 $f(x+h)$ 的泰勒展开式中 h 的各次幂的系数定义为函数 f 的各阶导数(微商)．至于不定积分，当然

可通过导数的逆运算来定义. 但是,这样一来,考虑的函数范围太窄了, 而且不用极限概念也无法讨论无穷级数的敛散性问题.所以,拉格朗日以幂级数为工具的代数方法,未能解决分析学的奠基问题.

到了19世纪,出现了一批杰出数学家, 他们积极为微积分学的重新奠基而努力. 首先要提到捷克的哲学教授布尔查诺(B.Bolzano).他开始将严格的论证引入数学分析学中. 1816年,他在二项展开公式的证明中,明确地提出了级数的收敛性概念,同时对极限、连续和变量也有较深入的理解.特别是他曾写出《无穷的悖论》一书,书中包含有许多真知灼见,可惜直到他去世后两年(1850年)才得出版.

分析学的重要奠基人, 公认是法国的多产数学家柯西(A. L. Cauchy). 他在1821~1923年间出版的《分析教程》和《无穷小计算讲义》, 堪称数学史上划时代的著作. 在那里, 他给出了分析学一系列基本概念的精确定义.例如,他给出了精确的极限定义,并把函数连续性、导数、微分、积分、无穷级数收敛性等重要概念,都建立在较坚实的基础上. 对此, 当年挪威杰出的青年数学家阿贝尔(N. H. Abel),在他1826年关于二项展开式的论文中曾如此赞扬了柯西的成就:"每一个在数学研究中喜欢严密性的人, 都应该读这本杰出的著作——《分析教程》."柯西的重要贡献, 就是抛弃了欧拉坚持的函数的显式表示以及拉格朗日的形式幂级数, 而引进了处理函数的新概念.

分析学的进一步严密化的工作，是由一位大器晚成的数学家外尔斯特拉斯(K. T. W. Weierstrass)作出的．这是他在1841～1856年担任中学教师时完成的工作．他的这些成果，直到1859年当他成为柏林大学的教师后，才逐渐为人们所获知．他首先改善了柯西关于极限概念的描述性定义，提出了免除直观的 $\varepsilon-\delta$ 和 $\varepsilon-N$ 陈述法．这便是今日数学分析教科书中普遍采用的极限表述法．往后，当讨论极限理论时，我们还将回到这个课题．

对于各种极限运算，首先要考虑极限是否总是存在？假如关于无理数还缺乏精确概念，而碰到变量的极限不是有理数时，则极限岂不就成为一个没有确切概念的未知对象？进一步，还可以提出"极限运算可否在实数系中畅行无阻"的问题．这些都归结为实数系的"完备性问题"．

首要的问题是无理数如何严格定义，以及如何证明它们也适合一般运算规律．于是，数学分析的最终奠基便归结到建立严格的实数系理论．这一历史任务是在柯西逝世后十余年，才由戴德金、康托尔(G. Cantor)、海涅(H. E. Heine)、外尔斯特拉斯等人不约而同地完成的．他们采取的途径各不相同，而有关著作竟在同一年（即1872年)发表，这也是数学史上的一个巧合．

数学分析发展到20世纪，已经像棵老树那样，产生了许多粗壮高大的新分支．特别是实变函数论(包括测度论与积分论)成为经典分析学的直接延伸．还应该提到的是，由阿伯拉罕·鲁滨生(Abrahm Robinson)在60年代

创始的"非标准分析学". 这是把久已废黜的莱布尼茨的无穷小分析法在崭新的逻辑基础上重新复活起来的新兴学科.现代数理逻辑学家非常重视这一巨大成就. 例如,不久前故世的权威人物哥德尔(K. Gödel)就曾说过:"不论从哪一方面看,非标准分析将会成为未来的数学分析."我们也同意这一观点,相信非标准分析的巨大功效,必将在今后充分显示出来.后面(本书第65页)将对这一学科的思想方法和基本结果作一简要介绍.

一、实数理论

关于实数的构造,已有三大派理论(也就是有三种构造方法),即戴德金的'分划',康托尔-海涅的'基本序列'和外尔斯特拉斯的'有界单调序列'. 这三种构造方法的共同点,都是利用有理数的某些集合来定义无理数. 当然,这是符合逻辑学上的'定义规则'的,即:凡是新概念,都必须根据已知概念来定义.

1. 从有理数集谈起

有理数集 Q 有一个重要性质,叫作'稠密性',是指:对于两个有理数 r_1、r_2,不管如何相接近(即 $|r_1-r_2|$ 不论如何小),在它们之间总有另一个有理数 r_3,例如取 $r_3=\frac{1}{2}(r_1+r_2)$ 即是. 当然,在 r_1、r_3 之间和 r_3、r_2 之间,还分别有另外的有理数. 如此类推,可知 r_1 与 r_2 之间存在

着无穷多个有理数.

在数轴上,以有理数为坐标的点(位置点),叫做有理点.于是,数集 Q 对应于数轴上的有理点集.既然点和数一一对应,即可用点代表数,用数表示点,故在数轴上谈论问题时,分析学上就常常习惯地将'点'和'数'两个名词混用不分.

Q 的稠密性,在数轴(坐标点组成的直线)上即表现为有理点集的稠密性.今后,也同样用Q表示有理点集.设 $r\in\mathbf{Q}$,则由稠密性可知:总存在有理点列 $r_n\in\mathbf{Q}$(其中 $n=1,2,3,\cdots$)以 r 为极限,即 $\lim\limits_{n\to\infty}r_n=r$,也即 $\lim\limits_{n\to\infty}|r-r_n|=0$.这表明有理点列之间的距离可以缩成零.由此可知,有理点与有理点之间不可能存在着大于零的'最小间距'.同样,也可以说,任何有理点 r 附近都不存在与 r 最近的而又异于 r 的有理点.

众所周知,单位正方形的对角线是长度为$\sqrt{2}$的线段,而$\sqrt{2}$并不是一个有理数.证明这一点是不难的,可用反证法:假设$\sqrt{2}=\dfrac{p}{q}$, $p\in\mathbf{Z}_+,q\in\mathbf{Z}_+$,并设 p 与 q 互素(即 p 与 q 不再存在大于1的公因子);两边取平方,可得$2q^2=p^2$,因此 p 为偶数,可记成 $p=2s$,而 $s\in\mathbf{Z}_+$;再代入原式,得 $2q^2=4s^2$,也即 $q^2=2s^2$,故 q 又必为偶数;于是,p 和 q 都有 2 的公因子.矛盾.

设想我们和古希腊的毕达哥拉斯 (Pythagoras) 一样,头脑中只有有理数概念,而根本不知道无理数为何物.这时,我们心目中的数轴(直线)上就只有坐标为有理

数的有理点.于是，以原点 O 为心，以长为$\sqrt{2}$的线段为半径画圆，在数轴上就截不出有坐标值的交点了．换言之，圆的弧线将在有理点集的'隙缝'中穿过去，因而根本得不到交点.

所以，如果不把有理点集间的隙缝填补起来，甚至连平面几何作图过程中"截取交点"一事有时都会遇到麻烦．更何况在分析学中有理点列的极限有时还可能不是有理点.因此，无理数的概念非引进不可.顺便提到，在古代，当毕达哥拉斯学派的门徒发现无理数$\sqrt{2}$时，曾引起了震动，造成了数学史上所谓的第一次危机.

2．戴德金分划

戴德金引出分划概念的想法，是从分析$\sqrt{2}$的定义方式得来的.他的原始思想也多少受到了欧几里得(Euclid)《几何原本》中关于'不可公度比'处理方法的启发.

由上段所论，已知$\sqrt{2}\notin \mathbf{Q}$．不难找到有理点序列$r_n<\sqrt{2}$和$r'_n>\sqrt{2}$(其中 $n=1,2,3,\cdots$)，使得根据$\mathbf{Q}$的稠密性能有$\lim\limits_{n\to\infty}(r'_n-r_n)=0$．于是，从直观上看，可以认为$\sqrt{2}$恰好成为两个有理点集$\{r_n\}$与$\{r'_n\}$的分划点(或'隙缝')．但是，我们尚不知无理数$\sqrt{2}$为何物，又如何能选取 r_n 使与$\sqrt{2}$作比较呢？这问题倒不难回答，因为我们总可以选取 $r_n\in\mathbf{Q}$，使 $r_n^2<2$．同理，可选取 $r'_n\in\mathbf{Q}$，使 $r'^2_n>2$.

满足上述条件的点集(序列)$\{r_n\}$与$\{r'_n\}$显然是非常多的.为了避免特殊性和不确定性，最好的办法就是直接利用点集 $\mathbf{Q}$ 本身作成的二分法.例如，我们可以把 $\mathbf{Q}$ 分

裂成‘左半集’A与‘右半集’A'：

$$A=\{r\mid r\in\mathbf{Q},\text{当 } r>0 \text{ 时 } r^2<2\},$$
$$A'=\{r'\mid r'\in\mathbf{Q},\text{当 } r'>0 \text{ 时 } r'^2>2\}.$$

当然，$A\cup A'=\mathbf{Q}$，且$A\cap A'=\varnothing$．这样一来，$\sqrt{2}$便对应于 A 与 A' 两者的分划点(分界处)，可记为$(A|A')$.

同理，无理数$\sqrt{3}$可以用分划$(B|B')$来定义，而 B 与B'的定义是

$$B=\{r\mid r\in\mathbf{Q},\text{当 } r>0 \text{ 时 } r^2<3\},$$
$$B'=\{r'\mid r'\in\mathbf{Q},\text{当 } r'>0 \text{ 时 } r'^2>3\}.$$

再例如，自然对数底 $e=\lim\limits_{n\to\infty}\left(1+\dfrac{1}{n}\right)^n=2.7182818\cdots$ 也是无理数，可用分划$(E|E')$来界定，其中

$$E=\left\{x\mid x\in\mathbf{Q},\ x\leqslant\left(1+\frac{1}{k}\right)^k,\ k\in\mathbf{Z}+\right\},$$
$$E'=\mathbf{Q}\setminus E.$$

这里仍然有 $E\cup E'=\mathbf{Q}$，$E\cap E'=\varnothing$.

综上所述，便可总结出普遍的分划概念：设 X 与 X' 为有理数集 $\mathbf{Q}$ 的两个非空子集，满足如下三个条件

(i)凡 X 中的有理数，均比 X' 中的有理数为小；

(ii) X 与 X' 的并集为 $\mathbf{Q}$，即 $\mathbf{Q}=X\cup X'$；

(iii) X 中无最大数，即：如果 $r\in X$，则必还有 $r'\in X$，而 $r<r'$.

这样，X 与 X' 便做成一个**戴德金分划**$(X|X')$，它代表一个**实数**.(或者干脆说，“分划就是实数”.)

注意上述条件 (iii)，只是要求左半集 X 恒无最大数，但右半集 $X'=\mathbf{Q}\setminus X$ 却可能包含最小数；那个最小

数当然也是有理数．此时，分划$(X|X')$就代表那个有理数．另一可能情况是，右半集X'也不含最小数．此时，分划$(X|X')$便是一个无理数．

这样，戴德金便完成了利用有理数来定义无理数的工作．

3．实数系的序结构

由上可知，每一分划$(X|X')$都由$\mathbf{Q}$的子集X唯一确定．因此，也可将分划简记为(X)，并称(X)为一实数．特别，当$X'=\mathbf{Q}\setminus X$不含最小数时，(X)便叫作无理数．

任意一对实数(X)和(Y)，如果$X\subseteq Y$，则规定大小顺序为$(X)\leqslant(Y)$．这也可记作$(Y)\geqslant(X)$．容易验证，实数间的顺序关系具有下列性质：

(i) 若$(X)\leqslant(Y)$、$(Y)\leqslant(Z)$，则$(X)\leqslant(Z)$；

(ii) 若$(X)\leqslant(Y)$且$(Y)\leqslant(X)$，则$(X)=(Y)$；

(iii) 对任意两个实数(X)和(Y)，必有$(X)\leqslant(Y)$或$(Y)\leqslant(X)$成立(也即两关系中必有一者成立)．

性质(i)、(ii)是显然的．现证性质(iii)：假如不等式$(X)\leqslant(Y)$不成立，则子集X不含于Y，故必有某元$r\in X$，但$r\notin Y$．于是，由分划性质(i)可知对任意$r'\in Y$恒有$r'<r\in Y'$．又因$r\in X$，故知$r'\in X$，这就表明$Y\subseteq X$，也即$(Y)\leqslant(X)$成立．

既然实数都具备顺序性质(i)、(ii)、(iii)，所以由一切实数构成的实数集$\mathscr{R}$是一个有序集．为了记法简便，今后也常用$x, y, z, t, \alpha, \beta, \gamma$等表示实数．

4．确界存在定理

确界存在定理是指：凡非空的有上界的实数子集合必有一个上确界(最小上界).

这是分析学中一条极其基本的重要定理．由它可以推导出一系列极为有用的等价命题．现在给出定理的证明如下：

设 $\mathscr{X}$ 是一个非空的有上界的实数集合，于是对 $\mathscr{X}$ 的任一上界(实数-分划)(B)，不等式关系

$$(X)\leqslant(B)$$

对一切$(X)\in\mathscr{X}$ 都成立．按实数的顺序定义，有

$$X\subseteq B.$$

现在我们取 $\mathscr{X}$ 中的一切 (X)所相应的集合 X 的并集：

$$B_0=\bigcup_{(X)\in\mathscr{X}}X.$$

在上述记法中，B_0 表示由所有 X 作成的并集，$\bigcup$ 表示'求并集'运算，$(X)\in\mathscr{X}$表示运算是对$\mathscr{X}$中的一切(X)(相应集合 X)进行的.

首先应证明集合 B_0 给出分划(B_0)．这需要验明分划条件(i)、(ii)、(iii)．自然，B_0 是个有理数集，B_0 与 $B_0'=\mathbf{Q}\setminus B_0$皆非空集．条件(ii)由 B_0'定义得知．为证(i)，任设 $r\in B_0$、$r'\in B_0'$．根据 B_0 定义可知必有 X_1 使 $r\in X_1$．于是必有 $r<r'$ (否则，$r'<r$ 时 $r'\in X_1\subseteq B_0$)．因此条件(i)恒满足．再由 $r\in B_0$ 可知 $r\in X_1\subseteq B_0$，因 (X_1) 为分划，故 X_1 中还有较大者 r_1 使得 $r<r_1$．所以 B_0 也满足条件(iii)．这就证明了 B_0 能给出实数(B_0).

其次再证明(B_0)是 $\mathscr{X}$ 的最小上界．因为对任意的$(X)\in\mathscr{X}$都有 $X\subseteq B_0$，即$(X)\leqslant(B_0)$，所以(B_0)是实数集

合$\mathscr{X}$的上界．又，对于$\mathscr{X}$的任意一个上界(B)，则$\mathscr{X}$中的一切(X)都合条件$(X)\leqslant(B)$，即$X\subseteq B$.因而$B_0\subseteq B$，故得$(B_0)\leqslant(B)$．这表明(B_0)为$\mathscr{X}$的上确界，通常记作$(B_0)=\sup\mathscr{X}$．确界存在定理得证．

作为定理的简单推论，可得命题："凡单调递增的有界数列必有极限．"事实上，数列作成一数集，它的上确界就是该数列的极限．（本书第20页讨论极限论时，还将回到这一命题．）同样地，可以证明"凡非空的有下界的实数集合$\mathscr{X}$必有下确界（最大下界），通常记作$\inf\mathscr{X}$．"

5．$\sqrt{2}\cdot\sqrt{3}=\sqrt{6}$ 的证明

用分划概念定义实数之后，还可在实数系上引进四则运算，并验明它们满足算术运算律（如交换律、结合律、分配律等），因此实数系成为一个数域．

先说一下加法运算．假设x、y分别是由分划(X)、(Y)给出的实数，即$x=(X)=(X|X')$，$y=(Y)=(Y|Y')$，当然X与Y都是有理数集.于是$x+y$也是一个实数，它由分划$(Z|Z')$给出，即

$$x+y:=(Z|Z'),$$

其中Z与Z'的定义是

$$Z=\{r\,|\,r=r_1+r_2,r_1\in X,\ r_2\in Y\},\quad Z'=\mathbf{Q}/Z.$$

由于在Z的定义里，$r=r_1+r_2=r_2+r_1$，故知$y+x$与$x+y$的定义分划相同，因此交换律$x+y=y+x$自然成立.不妨简记作$z=x+y$.

类似地，还可利用分划方法定义乘法运算$u=x\cdot y$.但需要考虑到负数乘负数得正数的问题，叙述方式上须

顾及一些技术细节，这里就只以$\sqrt{2}\cdot\sqrt{3}=\sqrt{6}$的证明为例，而不作普遍论述了.

在讨论戴德金分划时，我们已给出$\sqrt{2}$和$\sqrt{3}$两个数的分划定义，即$\sqrt{2}=(A|A')$，$\sqrt{3}=(B|B')$. 同理，可把$\sqrt{6}$表为$\sqrt{6}=(C|C')$，其中 C 与 C' 的定义是

$$C=\{r\mid r\in\mathbf{Q},\text{当 } r>0 \text{ 时 } r^2<6\},\qquad C'=\mathbf{Q}\setminus C.$$

$\sqrt{2}$与$\sqrt{3}$两个数的乘积$\sqrt{2}\cdot\sqrt{3}$可由下述分划给出：

$$\sqrt{2}\cdot\sqrt{3}=(A|A')\cdot(B|B')=(D|D'),$$

这里，有理数集 D 与 D' 的定义是

$$D=\{r\mid r\in\mathbf{Q},\text{当 } r>0 \text{ 时 } r \text{ 可表为 } r=r_1\cdot r_2,$$
$$\text{其中 } r_1^2<2,\ r_2^2<3,\ r_1、r_2\in\mathbf{Q}\},$$
$$D'=\mathbf{Q}\setminus D.$$

因此，为证明$\sqrt{2}\cdot\sqrt{3}=\sqrt{6}$，只须证明$(C|C')=(D|D')$，也即 $C=D$.

按 D 的定义，知 $r^2=(r_1\cdot r_2)^2=r_1^2\cdot r_2^2<2\cdot3=6$，故 $D\subseteq C$. 另一方面，设 $r\in C$，且 $r>0$，则 $r^2<6$. 于是，根据有理数集的稠密性，总可选取合乎条件 $r_1^2<2$、$r_2^2<3$ 的正有理数 r_1、r_2，使得 $r=r_1\cdot r_2\in D$. 从而推知 $C\subseteq D$. 这就证得了 $C=D$，也即$\sqrt{2}\cdot\sqrt{3}=\sqrt{6}$获证.

[附记] 在上述证明中，用到了有理数集 $\mathbf{Q}^2=\{r^2\mid r\in\mathbf{Q},\ r\geqslant0\}$的稠密性. 事实上，对任意两个有理数 r_1、r_2，不论$|r_1-r_2|$如何小，总有另一有理数，例如 $r^2=\left(\frac{r_1+r_2}{2}\right)^2$ 介于 r_1^2 与 r_2^2 之间，且在任意有界集上 $r_1^2-r_2^2=(r_1-r_2)(r_1+r_2)$的绝对值可随着 $|r_1-r_2|$的无限变

小而变小．今对任给 $r>0$，如果 $r^2<6$，则 $d=6-r^2>0$，且 $d<6$．于是，由稠密性可知存在有理数 $r_1>0$，使得 $2-\frac{1}{4}d\leqslant r_1^2<2$．从而，令 $r_2=\frac{r}{r_1}$，可得

$$r_2^2=\left(\frac{r}{r_1}\right)^2=\frac{r^2}{r_1^2}=\frac{6-d}{r_1^2}\leqslant\frac{6-d}{2-\frac{1}{4}d}<3.$$

这表明确实存在正有理数 r_1、r_2，满足条件：$r_1^2<2$、$r_2^2<3$，使得

$$r=r_1\cdot r_2.$$

6．实数系的戴德金性质

引进了由分划定义的实数之后，我们就有了确界存在定理．于是，实数系便具有连续性，即"戴德金性质"：如果把全体实数分成两个非空集合 A 和 B，使得 A 中每一个数都小于 B 中每一个数；那末，或者 A 中含最大数，或者 B 中含最小数，两者中恰有一者为真．

最近在"一条被漏掉了的基本定理"一文中（见《数理化信息》第二辑，辽宁教育出版社 1981 年版），作者张景中指出：正好比关于自然数有数学归纳法那样，关于实数可以有"连续归纳法"，即："假设 P_x 是关于实数 x 的命题．如果满足两个条件：

(i) 存在某实数 x_0，使对一切实数 $x<x_0$，有 P_x 成立；

(ii) 若对一切实数 $x<y$，有 P_x 成立，则有 $\delta_y>0$，使 P_x 对一切实数 $x<y+\delta_y$ 成立．

那末，命题 P_x 便对一切实数 x 都成立．" 这是一个

有趣的命题，而且用起来也很方便.不难发现这一命题和“戴德金性质”是等价的．读者如有兴趣，不妨加以验证.

限于篇幅，关于利用基本序列类和有界单调序列类(各个“等价类”)来界定实数的方法，这里就不作介绍了.事实上，这两种方法构成的实数理论和分划作成的理论是等价的.读者欲知其详，请参看《实数的构造理论》一书(王建午、曹之江、刘景麟编，人民教育出版社1981年版).

二、极　限　论

极限论要回答的主要问题有两个：一是有哪些判断极限存在性的准则？就是说，在什么样的条件下，极限才保证能存在？二是如何计算极限？就是问，极限算法满足哪些基本运算律？

1．关于序列极限

在柯西时代，极限概念是利用直观语言表述的.例如说什么“某序列中的数越来越接近某个数”，“两者之差可任意地减小”，“某变量可任意地增大”，等等．当然，这些话并没有说错，直至目前，一些入门性的教科书中还经常使用这些说法.但是，当需要分析稍复杂一点的极限过程时，比如说，想要证明下述论断†

---

† 为了表述简便，本书中采用下述记法：假设从某个命题 A(定理、论断或条件)可推断出另一命题 B，则记作 $A\Rightarrow B$．如果反过来还能从 B 推出 A，则可以记作 $A\Longleftrightarrow B$.这时就说 A 和 B 两个命题是等价的.

$$\lim_{n\to\infty} a_n = a \quad \Rightarrow \quad \lim_{n\to\infty} \frac{a_1 + a_2 + \cdots + a_n}{n} = a,$$

那就会感到直观性的极限表述法无能为力了.

下面我们先讨论序列极限的概念.最自然的办法,就是从定性表述法过渡到定量表述法，这也就是外尔斯特拉斯所建议的表述方式.

我们说一个无穷序列 $\{a_n\} = \{a_1, a_2, \cdots, a_n, \cdots\}$ 具有极限 a,是指:如果当 n 足够大时，差值 $a_n - a$ 可以任意地小.用记号来表示,就是:

$$n \to \infty \quad \Rightarrow \quad a_n \to a; \qquad \text{或写作} \quad \lim_{n\to\infty} a_n = a.$$

有时也可简记为 $\lim a_n = a$,并称 $\{a_n\}$ 收敛于 a.

关于上述极限的定量表述，叫做“ε-ω 准则”，即“$\lim\limits_{n\to\infty} a_n = a$ 的必要与充分条件是:对于每个 $\varepsilon > 0$,都存在一个数 $\omega \in \mathbf{Z}_+$,使得

$$n > \omega \quad \Rightarrow \quad |a_n - a| < \varepsilon.\text{”}$$

通常,ε 越小,ω 就需要越大. 当然，ω 依赖于 ε，故有时也记作 $\omega = \omega(\varepsilon)$.

注意:正因为 ε 具有任意性(也即可以任意地接近于0),所以不等式 $|a_n - a| < \varepsilon$ 表明 a_n 必将趋近于 a,只要当 $n \to \infty$ 时即可.

采用 ε-ω 准则，可知前面提到的那个较复杂的极限命题,可改述为：如果 $\lim\limits_{n\to\infty} a_n = a$,那末对任给 $\varepsilon > 0$,总存在 $\omega \in \mathbf{Z}_+$,使得

$$n > \omega \quad \Rightarrow \quad \left| \frac{a_1 + a_2 + \cdots + a_n}{n} - a \right| < \varepsilon.$$

现在我们来证明这命题．首先，由 $\lim\limits_{n\to\infty} a_n = \alpha$ 可知 $\{|a_n|\}$为有界数列，因此有上界 K，使得$|a_n| < K$(其中 $n\in \mathbf{Z}_+$)．从而 $|a_n-\alpha| \leqslant |a_n| + |\alpha| < K+|\alpha|$．

对任给 $\varepsilon>0$，总可取充分大 $m\in\mathbf{Z}_+$，使得

$$n\geqslant m \quad\Rightarrow\quad |a_n-\alpha| < \frac{\varepsilon}{2}.$$

对固定的 m，当然还可选取更大的 $\omega>m$，使得

$$n>\omega \quad\Rightarrow\quad \frac{m(K+|\alpha|)}{n} < \frac{\varepsilon}{2}.$$

于是，当 $n>\omega$ 时，便得出

$$\left|\frac{a_1+a_2+\cdots+a_n}{n}-\alpha\right| = \left|\frac{(a_1-\alpha)+(a_2-\alpha)+\cdots+(a_n-\alpha)}{n}\right|$$

$$\leqslant \frac{|a_1-\alpha|+\cdots+|a_m-\alpha|}{n} + \frac{|a_{m+1}-\alpha|+\cdots+|a_n-\alpha|}{n}$$

$$\leqslant \frac{m(K+|\alpha|)}{n} + \sup\{|a_{m+1}-\alpha|,\ \cdots,\ |a_n-\alpha|\}$$

$$< \frac{\varepsilon}{2}+\frac{\varepsilon}{2} = \varepsilon.$$

这就是我们所要证明的．

读者可以想一想，假如不采用 ε-ω 表述法，仅凭关于极限的直观性语言，将如何能说清楚上面的极限论断呢！

假设序列$\{a_n\}$和$\{b_n\}$分别有极限 $\lim a_n=\alpha$ 和 $\lim b_n = \beta$ (也即$\{a_n\}$ 与$\{b_n\}$分别收敛于 α 和 β)，则不难验证下列运算律：

1° 若 $\lambda\in\mathscr{R}$，则 $\lim(\lambda a_n)=\lambda\lim a_n$．

2° $\lim(a_n+b_n)=\lim a_n+\lim b_n$．

3° $\lim(a_n-b_n)=\lim a_n-\lim b_n$．

4° $\lim(a_n b_n)=(\lim a_n)(\lim b_n)$

5° 若 $b_n\neq 0$（其中 $n\in\mathbf{Z}_+$）且 $\lim b_n\neq 0$，则

$$\lim\frac{a_n}{b_n}=\frac{\lim a_n}{\lim b_n}.$$

6° 若 $a_n\leqslant b_n$（其中 $n\in\mathbf{Z}_+$），则 $\lim a_n\leqslant\lim b_n$.

7° 若 $\alpha=\beta$，且有序列 $\{c_n\}$ 满足条件 $a_n\leqslant c_n\leqslant b_n$（其中 $n\in\mathbf{Z}_+$），则 $\lim c_n=\alpha=\beta$.

上述诸结论的证明都十分简易，读者不难采用 ε-ω 表述法把精确的证明步骤一一写出来，这里从略.

2. 柯西收敛准则

如果给定一个序列$\{x_n\}$，将如何判断它有没有极限呢？对于这个问题，柯西首先找到了一个简便易用的准则，此即：序列$\{x_n\}$具有极限的必要与充分条件是

$$m、n\to\infty \quad\Rightarrow\quad |x_m-x_n|\to 0.$$

换句话说，序列 $\{x_n\}$收敛于某极限的充要条件是：对任给 $\varepsilon>0$（不论如何小），总存在 $\omega\in\mathbf{Z}_+$，使得

$$m、n>\omega \quad\Rightarrow\quad |x_m-x_n|<\varepsilon.$$

下面来证明这个重要命题（世称"柯西收敛性原理"）. 先来证明条件的必要性. 如果已知$\{x_n\}$有极限，比方说 $\lim x_n=\alpha$. 于是存在 $\omega\in\mathbf{Z}_+$，使得

$$m>\omega、n>\omega \quad\Rightarrow\quad |x_m-\alpha|<\frac{\varepsilon}{2}、\ |x_n-\alpha|<\frac{\varepsilon}{2}.$$

从而，

$$\begin{aligned}|x_m-x_n|&=|x_m-\alpha+\alpha-x_n|\\&\leqslant|x_m-\alpha|+|x_n-\alpha|\end{aligned}$$

$$<\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon.$$

这就表明 m、$n\to\infty\Rightarrow|x_m-x_n|\to 0$. 故知条件是必要的.

再来证明条件的充分性.假设条件已满足. 先取 $\varepsilon=1$,则有某数 $\omega_0\in\mathbf{Z}_+$,使得

$$m、n\geqslant\omega_0 \quad\Rightarrow\quad |x_m-x_n|<1,$$

于是有

$$|x_m|=|x_m-x_{\omega_0}+x_{\omega_0}|\leqslant|x_m-x_{\omega_0}|+|x_{\omega_0}|\leqslant 1+|x_{\omega_0}|$$

$(m\geqslant\omega_0)$. 由此可知$\{x_m|m\geqslant\omega_0\}$为有界数集. 因为$\{x_m|m<\omega_0\}$也是有界的,所以整个数列 $\{x_n\}$是有界的.

根据确界存在定理, $\{x_n\}$必有上确界. 不妨记它为 $\beta=\sup\{x_n\}$. 于是,在$(\beta-\varepsilon,\ \beta)$的区间内,总有 $\{x_n\}$中的点. 逐步缩小 ε, 便可找到一个子序列$\{x_{m_k}\}$, 使得 $x_{m_k}\to\beta$,也即 $\lim x_{m_k}=\beta$.

既设条件已满足,故有

$$m_k>\omega、n>\omega \quad\Rightarrow\quad |x_{m_k}-x_n|<\varepsilon.$$

注意上列不等式对一切 $m_k>\omega$ 都成立, 当然可令 $m_k\to\infty$,从而得 $n>\omega\Rightarrow|\beta-x_n|\leqslant\varepsilon$. 这正好表明

$$\lim x_n=\beta.$$

因此,在条件满足的情况下,$\{x_n\}$果然有极限. 这又证明了条件的充分性.命题证毕.

值得引起注意的是,柯西准则的重要性就在于,它只是根据数列自身的结构来判断极限的存在性. 也正因这一点,这准则具有广泛的应用价值.

今后, 我们把具有性质"m、$n\to\infty\Rightarrow|x_m-x_n|\to 0$" 的

序列$\{x_n\}$，一概叫做“基本序列”．于是，柯西准则即断言：“凡基本序列都有极限.” 或换句话说：“凡基本序列都是收敛的.”一般教科书中，也常常把基本序列干脆称为“柯西序列”．

3. 一元函数

下面我们要讨论函数极限概念．为此，先谈谈较一般的函数概念．

函数就是对应关系（或称映射关系）．例如$y=f(x)$，就表示将变元x按照法则f对应到y的关系，可以简记作$f:x\mapsto y$．要紧的是，需要说明x的活动范围，通常称之为函数f的“定义域”（或称“变域”）；还要指出y取值的范围，叫作函数f的“值域”．假设$X\subseteq\mathscr{R}$和$Y\subseteq\mathscr{R}$是两个非空的实数点集，那末

$$f:\quad X\mapsto Y$$

即表示以X为定义域和以Y为值域的实变函数．有时，也把f叫做将X映入Y的“映射”．因为在一般情形下，函数对应的所有的值未必填满Y，所以通常只是说f把X映入Y．

特别，如果$\{y\,|\,y=f(x),\ x\in X\}$就是$Y$，则称$f$把$X$映满$Y$．此时，就把函数$f$叫作从$X$到$Y$的“满射”．函数值的集合可以简记为$f(X)$，即$f(X)=\{y\,|\,y=f(x),\ x\in X\}=\{f(x)\,|\,x\in X\}$．因此，$f$成为满射的条件即$f(X)=Y$．

4. 函数极限概念

设$S\subseteq\mathscr{R}$为一无穷点集，$c\in\mathscr{R}$是一个点．如果对每

个 $\delta>0$,总存在点 $x\in S$,使得 $|x-c|<\delta$,则称点 c"依附于"S.

容易明白，点 $c\in\mathscr{R}$ 依附于 S 的必要与充分条件是:对任给 $\delta>0$,都有 $S\cap[c-\delta,\ c+\delta]\neq\varnothing$. 又由于 δ 可依次缩小,故充要条件还可改述为:S 中恒存在收敛于 c 的点列 x_n(注意,这里并不要求点 c 含于 S 中).

现在我们来讨论如下的函数极限概念:

$$x\in S、x\to c \quad\Rightarrow\quad f(x)\to\alpha.$$

这个概念的定量表述法，就是外尔斯特拉斯引入的"ε-δ 准则":

考虑函数 f: $S\mapsto\mathscr{R}$.假设点 $c\in\mathscr{R}$ 依附于 S. 如果对每个 $\varepsilon>0$,总有一个 $\delta>0$ (只依赖于 ε 和 f),使得

$$x\in S、|x-c|<\delta \quad\Rightarrow\quad |f(x)-\alpha|<\varepsilon,$$

则称 $f(x)$于 $x\to c$ 时有极限 α,可记作

$$\lim_{x\in S,x\to c} f(x)=\alpha.$$

有时也简记作 $f(x)\to\alpha$(当 $x\to c$, $x\in S$ 时).

必须注意,在上述定义里,一般并不要求 c 属于它所依附的点集 S. 当然,$c\in S$ 只是一种特殊情形,只在此特殊情形下 $f(c)$是有定义的.

类似于序列极限的情形，关于函数极限也有同样的运算律:设 $S\subseteq\mathscr{R}$ 为一个点集,$c\in\mathscr{R}$ 为依附于 S 的一个点.令 f: $S\mapsto\mathscr{R}$, g: $S\mapsto\mathscr{R}$,且

$$\lim_{x\to c} f(x)=\alpha,\quad \lim_{x\to c} g(x)=\beta,$$

那末,我们有如下的运算律:

2° $\lim\limits_{x\to c}(f+g)(x)=\lim\limits_{x\to c} f(x)+\lim\limits_{x\to c} g(x)$,

3° $\lim_{x\to c}(f-g)(x)=\lim_{x\to c}f(x)-\lim_{x\to c}g(x)$,

4° $\lim_{x\to c}(fg)(x)=(\lim_{x\to c}f(x))(\lim_{x\to c}g(x))$.

又，如果 $g(x)\neq 0\ (x\in S)$，且 $\beta\neq 0$，则

5° $\lim_{x\to c}\dfrac{f(x)}{g(x)}=\dfrac{\lim_{x\to c}f(x)}{\lim_{x\to c}g(x)}=\dfrac{\alpha}{\beta}$.

如果 $\alpha=\beta$ 且 ϕ: $S\mapsto\mathscr{R}$ 为满足下列条件的函数

$$f(x)\leqslant\phi(x)\leqslant g(x), \qquad (x\in S)$$

则 $\phi(x)$在 c 点有极限，且

7° $\lim_{x\to c}\phi(x)=\lim_{x\to c}f(x)=\lim_{x\to c}g(x)$.

读者当不难利用 $\varepsilon-\delta$ 准则去验证上列各个极限式. 又，对应于序列极限的运算律 1°和 6°，读者可自行补出.

最后，还值得注意的是，如果把 x_n 改记为$x_n=f(n)$，则数集 $\{x_n\}=\{x_n|n\in\mathbf{Z}_+\}$ 也可看成为函数 $f:\mathbf{Z}_+\mapsto\mathscr{R}$ 的值域. 但序列极限 $\lim_{n\to\infty}x_n=\lim_{n\to\infty}f(n)$却不能照搬上面函数极限的 ε–δ 表述方式，原因是∞并非$\mathscr{R}$中的一个点. 当然，也不可说它是依附于 $\mathbf{Z}_+$. 总之，必须认为 ε–ω 准则和 ε–δ 准则是两个不同的准则.

三、连续性概念

仅有关于极限的简单运算律，还不足以处理许多常见的函数极限计算问题. 例如，设 $f:\mathscr{R}\mapsto\mathscr{R}$，又设 $\lim_{n\to\infty}x_n=a$，人们会问下述计算法则

$$\lim_{n\to\infty}f(x_n)=f(\lim_{n\to\infty}x_n)$$

是否成立？又，例如 $g: \mathscr{R}\mapsto\mathscr{R}$，那末试问是否成立下列算法

$$\lim_{n\to\infty} f(g(x_n)) = f(\lim_{n\to\infty} g(x_n)) = f(g(\lim_{n\to\infty} x_n))?$$

这些问题都归结到函数 f、g 是否具有连续性的问题.

1. 连续性概念

设 $f: S\mapsto\mathscr{R}$. 如果对任给 $\varepsilon>0$，总可找到 $\delta>0$，使得当 $|x-y|<\delta$（其中 x, $y\in S$）时便保证有 $|f(x)-f(y)|<\varepsilon$，这样，就说 f 是定义域 S 上的**连续函数**. 也可以说，f 属于连续函数类 $\mathscr{L}(S)$，可记作 $f\in\mathscr{L}(S)$.

显然，上述概念也可简记作

$$|x-y|\to 0 \text{（其中} x\text{、}y\in S\text{）} \quad \Rightarrow \quad |f(x)-f(y)|\to 0.$$

［例］ 易证函数 $f(x)=\sqrt{x}$ 在区间[0,1]上连续. 事实上，对任给 $\varepsilon>0$，可决定 $\delta>0$，使当 $|x-y|<\delta$ 时有

$$|\sqrt{x}-\sqrt{y}|<\varepsilon, \qquad \text{其中 } x\text{、}y\in[0,\ 1].$$

为此目的，我们写出

$$|\sqrt{x}-\sqrt{y}| = |\sqrt{x}-\sqrt{y}|\cdot\frac{\sqrt{x}+\sqrt{y}}{\sqrt{x}+\sqrt{y}} = \frac{|x-y|}{\sqrt{x}+\sqrt{y}}.$$

注意到 $\sqrt{x}+\sqrt{y}\geqslant\sqrt{|x-y|}$（显然，由两边平方即可看出），由此易推知

$$|\sqrt{x}-\sqrt{y}| = \frac{|x-y|}{\sqrt{x}+\sqrt{y}} \leqslant \frac{|x-y|}{\sqrt{|x-y|}} = \sqrt{|x-y|}.$$

故只须选取 $\delta=\varepsilon^2$，便可得出所欲证明的事实：

$$|x-y|<\delta \quad \Rightarrow \quad |\sqrt{x}-\sqrt{y}|\leqslant\sqrt{|x-y|}<\delta^{\frac{1}{2}}=\varepsilon.$$

设 $f\in\mathscr{L}(S)$，而 $x\in S$，则由 f 的连续性可知

$$x_n\to x \text{（其中 } x_n\in S\text{）} \quad \Rightarrow \quad f(x_n)\to f(x).$$

换言之，$\lim\limits_{n\to\infty} f(x_n)=f(\lim\limits_{n\to\infty} x_n)=f(x)$.

2. 一点处的连续性

考虑 $f: S\mapsto\mathscr{R}, \alpha\in S$. 如果 $|x-\alpha|\to 0$(其中 $x\in S$)$\Rightarrow$ $|f(x)-f(\alpha)|\to 0$，则称 $f(x)$ 在 $x=\alpha$ 点处**连续**.此时，按照连续性定义，便有

$$\lim_{x\to\alpha} f(x)=f(\alpha).$$

特别，如果只在 $x>\alpha$，$x\to\alpha$ (其中 $x\in S$)时才有 $f(x)\to\beta$，则称 β 为函数在 α 点的**右侧连续**，可记作

$$\lim_{x\to\alpha+} f(x)=\beta=f(\alpha_+).$$

同样可以定义函数在 α 点的**左侧连续**

$$\lim_{x\to\alpha_-} f(x)=f(\alpha_-).$$

如果 $f(\alpha_-)=f(\alpha_+)=f(\alpha)$，则就表明函数 f 在 α 点处是连续的. 而当 $f(\alpha_+)\neq f(\alpha_-)$时，$|f(\alpha_+)-f(\alpha_-)|$ 就叫作函数在 α 点处的**跃度**.

3. 复合函数的连续性

按照设计工程师的观点，不妨把函数 $y=f(x)$: $X\mapsto Y$ 看成是一个黑盒子(一种复杂的电子仪器)，如下图：



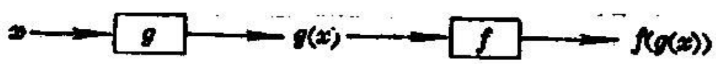
函数定义域 X 中的每一 x 都是一个可能的输入，而所有可能的输出 y 作成值域 $Y=f(X)=\{y\}$. 当然，这个观点也可推广到多元函数.

复合函数 $f(g(x))$的精确定义是：设 $g: S_1\mapsto S_2$，$f: S_2\mapsto S_3$，这里 S_1、S_2、S_3 均为 $\mathscr{R}$ 的子集(可以是区间，或 $\mathscr{R}$ 本身). 于是，用 $f\cdot g$ 表示的复合函数就是对定义域 S_1

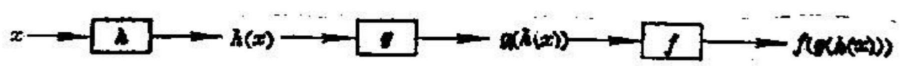
中的任何 x 其值为 $f(g(x))$ 的函数，也即

$$f\circ g: \quad S_1 \mapsto S_3.$$

可用图表示为



上述概念不难推广到三重复合函数，如 $f\{g[h(x)]\}$ 等，见下图：



由上图可以看出，三重复合函数既可解释成 $(f\circ g)\circ h$，又可理解为 $f\circ(g\circ h)$．因此，我们有

$$(f\circ g)\circ h = f\circ(g\circ h).$$

上述性质叫作复合函数的“可结合性”．

容易证明：两个连续函数的复合仍是连续的．例如，设 f: $\mathscr{R}\mapsto\mathscr{R}$ 与 g: $\mathscr{R}\mapsto\mathscr{R}$ 均为连续函数，则 $f\circ g(x)=f[g(x)]$，$f\circ g$: $\mathscr{R}\mapsto\mathscr{R}$．于是，由 g 的连续性知

$$|x-y|\to 0 \quad \Rightarrow \quad |g(x)-g(y)|\to 0,$$

又由 f 的连续性知

$$|g(x)-g(y)|\to 0 \quad \Rightarrow \quad |f[g(x)]-f[g(y)]|\to 0.$$

总之，有

$$|x-y|\to 0 \quad \Rightarrow \quad |f[g(x)]-f[g(y)]|\to 0.$$

这就是所要证明的结论，即 $f\circ g(x)\in\mathscr{C}(\mathscr{R})$．

4．三个基本命题

设 $f\in\mathscr{C}(I)$，$I=[a, b]$．则有如下三个看起来明显而证起来却很费口舌的命题：

最大值定理 f 必在区间 I 的某点 c 达到它的最大值．(当然，f 也会在另一点达到最小值．)

中间值定理 设 m 是 $f(a)$ 和 $f(b)$ 之间的任何数值，则在区间 I 内必有一点 c，使 $f(c)=m$.

反函数定理 设函数 $y=f(x)$ 是区间 I 上的单调上升函数(或单调下降函数)，那末它的反函数 $x=g(y)$ 必存在，而且是区间$[f(a), f(b)]$(或区间 $[f(b), f(a)]$)上的连续函数.

上述最大值定理和中间值定理的证明，要用到上确界存在定理和单调有界数列的极限存在性定理. 至于反函数定理，证明较为简单. 限于篇幅，这里就不去细说了.

5. 连续模数

在 $I=[a, b]$上的各种连续函数，有的升降变化很激烈，有的变化较平稳，那末，如何衡量这种变化的平稳性程度呢？"连续模数"就是对点距不超过一定数值（例如 δ)时函数值变差的一种度量. 令 $f\in\mathscr{C}(I)$， $I=[a, b]$. 则对任给 $\delta>0$，下述数量

$$\omega(\delta)=\omega(\delta; f)=\sup\{|f(x')-f(x'')| \mid |x'-x''|\leqslant\delta\}$$

便叫作函数 f 在 I 上的连续模数，这里的上确界 sup 是对一切可能的点 x'、$x''\in[a, b]$，$|x'-x''|\leqslant\delta$ 而取的. 对给定的 δ 而言，显然 $\omega(\delta)$越小，就表明 f 的变化越平稳. 特别， $\omega(\delta)=0$ 即表明 f 是一个常数. 读者当不难验明下列诸性质：

1° 函数 $\omega(\delta)$随着 δ 的变小而变小.

2° 对每一正整数 n，恒有 $\omega(n\delta)\leqslant n\omega(\delta)$. 事实上，这相当于下列不等式

$$\sup\{|f(x)-f(y)| \mid |x-y|\leqslant n\delta\}$$
$$\leqslant\sum_{i=1}^{n}\sup\{|f(x_i)-f(y_i)| \mid |x_i-y_i|\leqslant\delta\}.$$

3° 对任意正数λ，都有不等式$\omega(\lambda\delta)\leqslant(\lambda+1)\omega(\delta)$. 事实上，令$[\lambda]$表示$\lambda$的整数部分，则易见

$$\omega(\lambda\delta)\leqslant\omega(([\lambda]+1)\delta)\leqslant([\lambda]+1)\omega(\delta)\leqslant(\lambda+1)\omega(\delta).$$

6. 一致连续性定理

一致连续性定理是指："在闭区间$I=[a,\ b]$上的连连函数$f(x)$，必定具有性质$\lim\limits_{\delta\to0}\omega(\delta)=0$." 此处所说的性质，就叫做$f$的一致连续性；也即对每个$\varepsilon>0$，都可找到一个$\delta>0$，使当$|x'-x''|\leqslant\delta$时，恒有$|f(x')-f(x'')|<\varepsilon$，其中$x'$、$x''\in[a,\ b]$.

证明 由$\omega(\delta)$的单调性，不妨取$\delta=\dfrac{1}{n}$（其中$n\in\mathbf{Z}_+$）. 于是，一致连续性条件也可改述为

$$\lim_{n\to\infty}\omega\left(\frac{1}{n}\right)=0.$$

今用反证法，假设$\lim\omega\left(\dfrac{1}{n}\right)=\varepsilon_0>0$. 则对每个$\dfrac{1}{n}$，都存在$x'_n$、$x''_n\in[a,\ b]$，$|x'_n-x''_n|\leqslant\dfrac{1}{n}$，使得

$$|f(x'_n)-f(x''_n)|\geqslant\varepsilon_0,\ (n\in\mathbf{Z}_+).$$

因为$\{x'_n\}$和$\{x''_n\}$为有界点集，故应用确界存在性定理，可知有子序列$x'_{n_k}\to x_0\in I$. 相应地，由$|x'_{n_k}-x''_{n_k}|\to0$可知又有$x''_{n_k}\to x_0$. 因此，由连续性即得

$$\lim_{k\to\infty}|f(x'_{n_k})-f(x''_{n_k})|=|f(x_0)-f(x_0)|=0.$$

这和原来假设的推论$|f(x'_{n_k})-f(x''_{n_k})|\geqslant\varepsilon_0$矛盾，这就证

明了定理结论 $\lim\limits_{n\to\infty}\omega\left(\frac{1}{n}\right)=0$.

〔反例〕 定义在开区间上的连续函数未必有一致连续性. 比如, 试考虑函数 $y=\frac{1}{x}$, $x\in(0,1)$. 则当 $x''-x'=\delta>0$(其中 $0<x'<x''\leqslant 1$), 可知

$$\left|\frac{1}{x'}-\frac{1}{x''}\right|=\left|\frac{\delta}{x'(x'+\delta)}\right|\to\infty \quad (\text{当 } x'\to 0+\text{时}).$$

因此, 对任给 $\delta>0$, 在 $(0,1)$ 上恒有

$$\omega(\delta)=\sup\left\{\left|\frac{1}{x'}-\frac{1}{x''}\right| \;\middle|\; |x'-x''|\leqslant\delta\right\}=\infty.$$

这表明, 函数 $\frac{1}{x}$ 在 $(0,1)$ 上不可能一致连续.

最后还要指出, 只须应用 ε–δ 准则即不难验证分析学中常用的各种初等函数(如多项式、三角函数、指数函数、对数函数以及它们的复合函数等等)都是各个定义域内的连续函数, 且是闭域上的一致连续函数. 当然, 对它们施加极限运算可以畅行无阻.

四、微积分主题浅释

1. 主题简释

什么是微积分的主题? 这在初等微积分教科书中已有明确回答: 微分学主要是处理函数变量(应变量)的变化率问题, 即讨论微商(导数)的计算法则和有关问题. 积

分学是处理微分学的反问题，即如何从变化率去寻求(包括分析、计算)原函数的问题．

在应用上，利用变化率来描写的数量是多不胜举的．例如曲线的斜率、变速运动的速度、交流电的电流强度、空间温度场的梯度以及现代经济学上的边际劳动生产率、边际税率等等，都是变化率(也即导数)概念的具体例子．反过来，从已知的曲线斜率、物体运动速度等变量，去寻求曲线所满足的方程和表示运动规律的函数或方程等，这一计算过程就用到了积分．

微积分是一门变量数学．它是通过合理的抽象模式(Abstract pattern)来表现变量间的种种普遍关系结构的．因此，比如当我们论述一般函数的导数与积分时，就只考虑它们间的纯粹的普遍形式关系．当然，这些形式关系(有时用公式或定理来表述)可以应用于各种合适的具体场合，被赋予种种具体解释，从而帮助解决实际问题．在初等微积分中，我们已经见到这门数学在几何学与经典力学领域中的大量应用实例．读者还可以从现代各门正在发展着的应用科学的庞大领域中，不断发现近、现代分析数学扮演着何等重要的工具角色！当然，一切应用都无非是普遍模式的特殊化和具体化而已．

人们对微积分主题的认识，是逐步深化的．在柯西时代，人们主要考虑连续函数的积分．在微分学中只讨论具有以连续函数为导数的可微函数类，以及可以展成幂级数的解析函数类．到了19世纪下半叶，不仅多元微积分学有了很大的发展，而且柯西的积分概念被拓广成黎曼(G.

F. B. Riemann）的积分概念，使得可以被积分的函数类大大超出了连续函数类．接着又由斯蒂尔吉斯（Th. J. Stieltjes）提出了比黎曼积分更广的积分概念，即现代应用数学中常用的“黎曼-斯蒂尔吉斯积分”．

自从19世纪末期出现了集合论与点集论之后，通过波莱尔（E. F. E. J. Borel）和勒贝格（H. L. Lebesgue）等人的相继努力，于20世纪开始的年代里又形成了勒贝格测度论与积分论．后来，在积分论的研究发展中，又出现了勒贝格-斯蒂尔吉斯积分、帕梯斯（Pittis）积分、波赫纳（S. Bochner）积分等概念．这些都是20世纪中叶前的成就．

微积分学的“灵魂核心”就是著名的牛顿-莱布尼茨公式，又叫做“积分学基本定理”．它表明了积分与微分（求微商）互为反运算过程的基本关系．因此，伴随着积分概念的每一次拓广，人们都要考察微商（导数）概念的相应更新，以便使上述基本定理采取新的合理形式．

古典的导数概念首先是由狄尼（U. Dini）拓广的．这样，许多不甚光滑的甚至具有激烈振荡性的函数，也可以具有狄尼导数．与此有关的一些发展，这里不能细说了．但应该指出，正因为微积分中的一些基本概念都已成功地拓广到更广泛的空间和更一般的函数类上，所以在现代的泛函分析中还出现了弗雷谢（R.M. Frechet）微分、索伯列夫（С.Л. Соболев）导数以及一般广义函数的导函数等概念．尽管这些概念的抽象程度极高，但都可以从古典的微分与导数概念中找到它们的思想本源．

2．柯西的积分概念

参照柯朗(R. Courant)的观点，先讲积分后讲微分也是个好办法.现在先来回顾柯西的积分定义.

设 $f\in\mathscr{C}(I)$, $I=[a, b]$. 记 $\Delta x=\dfrac{b-a}{n}$, $n\in\mathbf{Z}_+$. 作区间 I 的等距划分

$$\Delta_n:\qquad a=x_0<x_1<x_2<\cdots<x_n=b,$$

这里 $x_j=a+j\Delta x$(其中 $j=0, 1, \cdots, n$). 令 $I_j=[x_{j-1}, x_j]$(其中 $1\leqslant j\leqslant n$)表示 I 的子区间.今在各个 I_j 中任取一点 $\xi_j\in I_j$. 作如下的求积和

$$S_n=\sum_{j=1}^{n}f(\xi_j)\Delta x.$$

可以证明,当 $n\to\infty$时,极限 $\lim\limits_{n\to\infty}S_n$ 必存在. 这个极限就叫作函数 $f(x)$在 $[a, b]$ 上的**定积分**,记作

$$\int_a^b f(x)dx\ :=\ \lim_{n\to\infty}S_n.$$

证明 需要证明的是,不论点 $\xi_j\in I_j$ 是怎样取法,极限总是唯一确定的.为此,令

$$M_j=\sup\{f(x)\,|\,x\in I_j\},\qquad m_j=\inf\{f(x)\,|\,x\in I_j\}.$$

分别作大和与小和:

$$\overline{S}_n=\sum_{j=1}^{n}M_j\Delta x,\qquad \underline{S}_n=\sum_{j=1}^{n}m_j\Delta x.$$

显而易见, $\underline{S}_n\leqslant S_n\leqslant\overline{S}_n$. 所以只须证明 $\lim\overline{S}_n=\lim\underline{S}_n$ 就可以了. 根据 f 在闭区间 I 上的一致连续性,可知对任给 $\varepsilon>0$,总存在 $\delta>0$, 使当 $\Delta x=\dfrac{b-a}{n}<\delta$ 时恒有$(M_j-m_j)<\varepsilon$(其中$j=1,\cdots,n$).由此可见,只须当$n>\dfrac{b-a}{\delta}$时,便有

$$\bar{S}_n - \underline{S}_n = \sum_{j=1}^{n} (M_j - m_j)\Delta x < \varepsilon \sum_{j=1}^{n} \Delta x = (b-a)\cdot\varepsilon.$$

注意，$\varepsilon>0$ 具有任意性，可使 $\varepsilon \to 0+$，故得

$$\lim_{n\to\infty}(\bar{S}_n - \underline{S}_n) = 0. \tag{1}$$

最后还须证明单个极限 $\lim \bar{S}_n$ 或 $\lim \underline{S}_n$ 是存在的．这样才能断言 $\lim \bar{S}_n = \lim \underline{S}_n = \lim S_n$．

考虑自然数子序列 2^m，其中 $m\in \mathbf{Z}_+$．则 $\Delta_{2^{m+1}}$ 为 Δ_{2^m} 的加细划分．故知 $\{\underline{S}_{2^m}\}$ 为单调增序列，$\{\bar{S}_{2^m}\}$ 为降序列，它们都是有界数列．按确界存在定理，可知它们都有确定极限值．又由(1)得知其极限是相同的．故可记作

$$\lim_{m\to\infty}\underline{S}_{2^m} = \lim_{m\to\infty}\bar{S}_{2^m} = L. \tag{2}$$

对于任意划分 Δ_n 和 Δ_{2^m}，考虑它们共同的加细划分 $\Delta_{n\times 2^m}$．显然有

$$\underline{S}_n \leqslant \underline{S}_{n\times 2^m} \leqslant \bar{S}_{n\times 2^m} \leqslant \bar{S}_n, \tag{3}$$

$$\underline{S}_{2^m} \leqslant \underline{S}_{n\times 2^m} \leqslant \bar{S}_{n\times 2^m} \leqslant \bar{S}_{2^m}. \tag{4}$$

今在(4)式中令 $m\to\infty$ 取极限，则由(2)得知

$$\lim_{m\to\infty}\underline{S}_{n\times 2^m} = \lim_{m\to\infty}\bar{S}_{n\times 2^m} = L.$$

从而，由(3)又得知 $\underline{S}_n \leqslant L \leqslant \bar{S}_n$．因此，根据(1)，便得到 $\lim\limits_{n\to\infty}\bar{S}_n = \lim\limits_{n\to\infty}\underline{S}_n = L$．这就完成了关于极限 $\lim\limits_{n\to\infty} S_n$ 存在唯一性的证明．

［特例］　试取 $\xi_j = \xi_j$，则有

$$\lim_{n\to\infty}\sum_{j=1}^{n} f\left(a + \frac{j}{n}(b-a)\right)\frac{b-a}{n} = \int_a^b f(x)dx.$$

由于 $\dfrac{b-a}{n} \to 0$（当 $n\to\infty$ 时），故人们常常把上式左端看成为无穷小量的求和．

3．积分中值定理

注意，在闭区间$[a, b]$上的连续函数$f(x)$都有最大值和最小值．也即存在M和m，使$m\leqslant f(x)\leqslant M$(其中$a\leqslant x\leqslant b$)．由求积和定义，可知

$$m\cdot(b-a)\leqslant S_n\leqslant M\cdot(b-a).$$

因此，

$$m\cdot(b-a)\leqslant\int_a^b f(x)dx\leqslant M\cdot(b-a).$$

这表明积分值介于$m(b-a)$与$M(b-a)$之间．自然，可记

$$\int_a^b f(x)dx=\mu\cdot(b-a),\ (\text{其中 } m\leqslant\mu\leqslant M).$$

由连续函数中间值定理，可知存在ξ(其中$a\leqslant\xi\leqslant b$)，使得$\mu=f(\xi)$．这就得到了'积分中值公式'（又称中值定理）：

$$\int_a^b f(x)dx=f(\xi)\cdot(b-a)\ (\text{其中 } a\leqslant\xi\leqslant b).$$

4．积分学基本定理

积分学基本定理是指，求微商(简称'求导')是积分的逆运算．设$f\in\mathscr{C}(I)$，$I=[\alpha, \beta]$而$[a, b]\subset[\alpha, \beta]$．则$[a,b]$上的积分$\int_a^b f(x)dx$依赖于积分的上限$b$和下限$a$．如取定下限$a$，而只让$b$变化，则积分值便是上限$b$的函数．改记$b$为$t$，则积分便成为$t$的函数，可记为

$$F(t)=\int_a^t f(x)dx. \tag{1}$$

现在令t获得一个增量Δt（仍设$t+\Delta t\in I$)，则

$$F(t+\Delta t)=\int_a^{t+\Delta t} f(x)dx. \tag{2}$$

由(2)减去(1)，得

$$F(t+\Delta t)-F(t)=\int_t^{t+\Delta t} f(x)dx. \tag{3}$$

应用积分中值定理，可得

$$F(t+\Delta t)-F(t)=f(\xi)\cdot\Delta t,\quad t\leqslant\xi\leqslant\Delta t.$$

于是，取极限，可求得 F 的导数为

$$F'(t)=\lim_{\Delta t\to 0}\frac{F(t+\Delta t)-F(t)}{\Delta t}=\lim_{\Delta t\to 0} f(\xi)=f(t).$$

这里，我们用到了 f 在点 t 的连续性.

综上所述，我们证得了由(1)式可导出

$$\frac{d}{dt}F(t)=f(t). \tag{4}$$

也即

$$\frac{d}{dt}\int_a^t f(x)\,dx=f(t). \tag{4}^*$$

这就表明$\frac{d}{dt}$确实是$\int$的逆运算.

通常，我们把由(1)式定义的函数 $F(t)$称为"不定积分". 因此，(4)*或(4)即表明 f 的不定积分经求导后便可还原成 f.

命题"(1)⇒(4)"叫作"积分学基本定理的第一形式". 利用这一形式，便可计算(1)式右端的积分. 人们自然想到应该找一个 $F(t)$，使之满足形如(4)式的简单微分方程. 但满足方程(4)的 $F(t)$并不是唯一的. 因为令 $F(t)$加上任意常数后仍然满足原来方程. 现在假定 $F(t)$是满足

方程(4)的任一函数，它未必恰好和下列函数

$$G(t)=\int_a^t f(x)dx \tag{5}$$

相一致. 这时，根据(4)*，我们仍然有$\frac{d}{dt}G(t)=f(t)$. 于是，和(4)式相比较，便得到

$$\frac{d}{dt}[G(t)-F(t)]=f(t)-f(t)=0 \quad (\text{其中 } t\in I).$$

因此 $G(t)-F(t)=K$(常数)，也即 $G(t)=F(t)+K$. 但由(5)式，可见 $G(a)=F(a)+K=0$，所以 $K=-F(a)$，也即 $G(t)=F(t)-F(a)$.

综上所述，我们证得了由(4)式可导出

$$\int_a^t f(x)dx=F(t)-F(a). \tag{6}$$

这样，我们得到了基本定理的第二形式，即命题"(4)⇒(6)". 当然，(6)式也可以表述成

$$\int_a^t \frac{d}{dx}F(x)dx=F(t)-F(a). \tag{6*}$$

如果有了微分概念，上式还可改写成

$$\int_a^t dF(x)=F(t)-F(a).$$

这又进一步表明积分是微分运算的逆运算.

上述互逆运算关系的发现，是牛顿、莱布尼茨的巨大功绩，所以理应把微积分学的创始归功于他们两人.

5. 微分中值定理

通过积分学基本定理，可把积分中值公式立即转化为微分中值公式. 事实上，我们有

$$\frac{F(b)-F(a)}{b-a}=\frac{1}{b-a}\int_a^b f(x)dx=f(\xi)=F'(\xi),$$

这里假定 F 在 I 上有连续导数，而 $a\leqslant\xi\leqslant b$.

无论是微分中值公式或积分中值公式，都只有当 a 与 b 相差较小时才成为估算的有用工具.

最后，让我们补充指出，在定积分的概念中，即使把等距分划改为一般的不等距分划，只须要求分划模数 $\|\Delta_n\|=\sup\{(x_j-x_{j-1})\mid 1\leqslant j\leqslant n\}\to 0$（当 $n\to\infty$ 时），仍可证得极限 $\lim\limits_{n\to\infty}S_n$ 的存在唯一性. 以后我们论述黎曼积分时，还将回到这一课题.

五、微商与微分

1. 微分与导数的关系

什么是微分？它和导数（微商）是什么关系？两者在概念上有无从属关系？这些都是微分学必须讨论的问题.

设函数 $y=f(x)$ 在点 x 有导数. 此时，下列极限存在：

$$y'=f'(x)=\lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x}=\lim_{\Delta x\to 0}\frac{f(x+\Delta x)-f(x)}{\Delta x}.$$

可以记作

$$\frac{\Delta y}{\Delta x}-y'=\alpha\to 0\quad（当\ \Delta x\to 0\ 时）.$$

也可改写成

$$\Delta y = y'\Delta x + \alpha\cdot\Delta x.$$

这里出现的余项 $\alpha\cdot\Delta x$,比 $\Delta x\to 0$ 趋零更快，故称为‘高级无穷小量’,一般可记作 $o(\Delta x)$. 又，趋零变量 α 本身也可记作 $o(1)$. 意思是 α 与 1 之比以零为极限.

注意,在关系式 $\Delta y = y'\Delta x + o(\Delta x)$中，$y' = f(x)$ 是个与确定的 x 有关的数值,当$\Delta x\to 0$ 时它是保持不变的. 换言之,Δy 的主要部分即 $y'\Delta x$ 与 Δx 始终保持线性比例关系,其比值就是y'.

反之,如果在已知点 x,函数 $y = f(x)$ 的改变量可以表成 $\Delta y = \lambda\cdot\Delta x + o(\Delta x)$,其中 λ 不依赖于 Δx，则称函数 f 在点 x 是‘可微的’,并且有导数 $y' = f'(x) = \lambda$.

我们把函数的改变量 Δy 的‘主要线性部分’记作$dy = \lambda\Delta x = y'\Delta x$，这就叫作 y 的‘微分’. 特别，如果取 $y = f(x) = x$,则 $y' = 1$,此时 $dx = 1\cdot\Delta x$. 因此，在任何函数 $y = f(x)$的微分表达式中,可记

$$dy = y'dx = y'\Delta x.$$

从而函数 y 的微分与自变量微分之比即可表成

$$y' = \frac{dy}{dx}.$$

不仅如此,如果 x 不是自变量， 而是一个新的自变量 t 的可微函数时,如 $x = \varphi(t)$,则由复合函数的求导法则,可得

$$dy = (f\circ\varphi(t))'dt = f'(x)\varphi'(t)dt = f'(x)\cdot dx.$$

这就表明,不管 x 是自变量或应变量(函数),y 的微分形式 $dy = f'(x)dx = y'dx$ 总是不变的. 这种形式上的不变性,在实际运算时是十分方便的.

综上所述，就一元函数 $y=f(x)$ 而言，可微的概念与可导的概念是一致的，而微分 dy 就是函数改变量 Δy 的线性主要部分．它和导数 y' 有关，但一般说来，和作为变化率的导数概念并无从属关系．当我们考虑多元函数的可微映射时，就可看出可微是比可导更强的概念．因为多元函数的可导性只须考虑沿着某一方向的导数是否存在，而可微性是要在一点的整个邻域中考虑的．

2．多元函数

在讨论一元函数时，已经指出：函数就是映射关系．这个观点可用以表述多元函数．我们假定读者已具备线代数的初步知识，已经知道数值矩阵作为线性映射，能把向量（点）变换成另一个向量（点）．

记 $\boldsymbol{x}=(x_1, \cdots, x_n)$，$\boldsymbol{y}=(y_1, \cdots, y_n)$，$\bar{\boldsymbol{x}}=(\bar{x}_1, \cdots, \bar{x}_n)$ 等为 n 维向量空间 $\mathscr{R}^n$ 中的列向量．在作运算时，理应写作

$$\boldsymbol{x}=\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \qquad \boldsymbol{y}=\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \qquad \bar{\boldsymbol{x}}=\begin{pmatrix} \bar{x}_1 \\ \vdots \\ \bar{x}_n \end{pmatrix}, \text{等}.$$

如果把向量中的诸分量看成坐标值，则 $\boldsymbol{x}$、$\boldsymbol{y}$ 等也就是空间 $\mathscr{R}^n$ 中的点．所以，在下面的论述中，点与向量二词可以交替使用．特别，$\mathscr{R}^n$ 作为向量空间，有如下一组基（单位向量）：

$$\boldsymbol{e}_1=\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \qquad \boldsymbol{e}_2=\begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \qquad \cdots, \qquad \boldsymbol{e}_n=\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

于是，每个向量（点）$\boldsymbol{x}\in\mathscr{R}^n$ 都可表示成基向量的线性组

合 $\boldsymbol{x}=x_1e_1+\cdots+x_ne_n$. 每两点 $\boldsymbol{x}$、$\boldsymbol{y}$ 间的内积，可写成 $\langle\boldsymbol{x}, \boldsymbol{y}\rangle=\sum_{i=1}^{n}x_iy_i$. 特别，$\langle\boldsymbol{e}_i, \boldsymbol{e}_i\rangle=1$，$\langle\boldsymbol{e}_i, \boldsymbol{e}_j\rangle=0$(其中 $i\neq j$). 通常实数的绝对值概念，可推广成"范数"概念，即

$$|\boldsymbol{x}|=\sqrt{\langle\boldsymbol{x},\boldsymbol{x}\rangle}=\left(\sum_{i=1}^{n}x_i^2\right)^{1/2}.$$

由线代数知，每个 $m\times n$ 的数值矩阵 $\boldsymbol{A}=(a_{ij})$(其中 $i=1,\cdots,m$; $j=1,\cdots,n$)都表现为一个从 n 维空间 $\mathscr{R}^n$ 到 m 维空间 $\mathscr{R}^m$ 的线性映射(变换). 它把每一点 $\boldsymbol{x}\in\mathscr{R}^n$ 映射成 $\mathscr{R}^m$ 中的一点 $\bar{\boldsymbol{x}}$，即

$$\boldsymbol{A}\boldsymbol{x}=\begin{pmatrix}a_{11}&\cdots&a_{1n}\\a_{21}&\cdots&a_{2n}\\\cdots&\cdots&\cdots\\a_{m1}&\cdots&a_{mn}\end{pmatrix}\cdot\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}=\begin{pmatrix}\bar{x}_1\\\bar{x}_2\\\vdots\\\bar{x}_m\end{pmatrix}=\bar{\boldsymbol{x}}\in\mathscr{R}^m.$$

其中 $\bar{x}_i=\langle a_i, \boldsymbol{x}\rangle=\sum_{j=1}^{n}a_{ij}x_j$, $a_i=(a_{i1}, \cdots, a_{in})\in\mathscr{R}^n$. 如果 $\boldsymbol{A}$ 的元素依赖于向量 $\boldsymbol{x}$，则记 $\boldsymbol{A}=\boldsymbol{A}_x$.

一般情形，令 $G\subset\mathscr{R}^n$ 为 $\mathscr{R}^n$ 的子集(可以是开区域或闭区域)，则可定义从 G 到 $\mathscr{R}^m$ 上的一般非线性映射(也即从 n 元到 m 元的函数)：

$$f:\ G\mapsto\mathscr{R}^m.$$

详细写出来，就是 $\bar{\boldsymbol{x}}=f(\boldsymbol{x})$, $\boldsymbol{x}\in G$, $\bar{\boldsymbol{x}}\in\mathscr{R}^m$，也即

$$\begin{aligned}f(x_1,\cdots,x_n)&=(f_1(x_1,\cdots,x_n),\cdots,f_m(x_1,\cdots,x_n))\\&=(\bar{x}_1,\cdots,\bar{x}_m)\in\mathscr{R}^m,\end{aligned}$$

其中 $\bar{x}_i=f_i(x_1, \cdots, x_n)$, $i=1,\cdots,m$. 因此，映射 $f:G\mapsto\mathscr{R}^m$ 包含有 m 个联立的从 G 到 $\mathscr{R}$ 的分量映射

$$f_i:\ G\mapsto\mathscr{R}.$$

这里假定 $\mathscr{R}^m$ 与 $\mathscr{R}^n$ 都具有按内积定义的范数.

3. 连续映射

对于 $f: G\mapsto\mathscr{R}^m$,如果对任给 $\varepsilon>0$,总存在 $\delta>0$,使当 $|\boldsymbol{x}-\boldsymbol{y}|<\delta(\boldsymbol{x}、\boldsymbol{y}\in G)$时,就保证有$|f(x)-f(y)|<\varepsilon$,则说 f 在点 $\boldsymbol{x}$ "连续".

显而易见,线性映射$\boldsymbol{A}$是连续的,因为有

$$|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{A}\boldsymbol{y}|=|\boldsymbol{A}(\boldsymbol{x}-\boldsymbol{y})|\to 0(\text{当}|\boldsymbol{x}-\boldsymbol{y}|\to 0\text{ 时}).$$

4. 可微映射

映射 $f: G\mapsto\mathscr{R}^m$ 称为在点$\boldsymbol{x}\in G$ 是"可微"的,如果存在一个线性映射$A_x: G\mapsto\mathscr{R}^m$,使对任意 $\boldsymbol{y}\in G$,

$$f(\boldsymbol{y})-f(\boldsymbol{x})=A_x(\boldsymbol{y}-\boldsymbol{x})+R(\boldsymbol{x},\ \boldsymbol{y}),$$

其中 $R(\boldsymbol{x},\ \boldsymbol{y})\in\mathscr{R}^m$, 而且 $|R(\boldsymbol{x},\ \boldsymbol{y})|=o(|\boldsymbol{y}-\boldsymbol{x}|)$, 也即

$$\lim_{|\boldsymbol{y}-\boldsymbol{x}|\to 0}\frac{|R(\boldsymbol{x},\boldsymbol{y})|}{|\boldsymbol{y}-\boldsymbol{x}|}=0.$$

映射 A_x 就叫作 f 在点 $\boldsymbol{x}$ 的"微分",可记作 $Df(\boldsymbol{x})$. 所以微分就是线性映射, 由它给出的 $[Df(\boldsymbol{x})](\boldsymbol{y}-\boldsymbol{x})$恰好就是改变量 $f(\boldsymbol{y})-f(\boldsymbol{x})$ (m 维向量) 的主要线性部分.

不难证明,如果 f 在点 $\boldsymbol{x}$ 可微,则对任意 $\boldsymbol{y}\in G$,

$$[Df(\boldsymbol{x})](\boldsymbol{y})=\lim_{\alpha\to 0}\frac{f(\boldsymbol{x}+\alpha\boldsymbol{y})-f(\boldsymbol{x})}{\alpha}.$$

证明 根据微分定义

$$f(\boldsymbol{y})-f(\boldsymbol{x})=[Df(\boldsymbol{x})](\boldsymbol{y}-\boldsymbol{x})+R(\boldsymbol{x},\ \boldsymbol{y}),$$

其中 $|\boldsymbol{y}-\boldsymbol{x}|\to 0\Rightarrow|R(\boldsymbol{x},\ \boldsymbol{y})|\to 0$.现在我们有

$$\begin{aligned}\frac{f(\boldsymbol{x}+\alpha\boldsymbol{y})-f(\boldsymbol{x})}{\alpha}&=\frac{1}{\alpha}[Df(\boldsymbol{x})](\alpha\boldsymbol{y})+\frac{1}{\alpha}R(\boldsymbol{x},\boldsymbol{x}+\alpha\boldsymbol{y})\\&=[Df(\boldsymbol{x})](\boldsymbol{y})+\frac{1}{\alpha}R(\boldsymbol{x},\boldsymbol{x}+\alpha\boldsymbol{y}),\end{aligned}$$

并且

$$\lim_{\alpha\to 0}\frac{|R(\boldsymbol{x},\boldsymbol{x}+\alpha\boldsymbol{y})|}{|\alpha|}=\lim_{\alpha\to 0}\frac{|\boldsymbol{y}|\,|R(\boldsymbol{x},\boldsymbol{x}+\alpha\boldsymbol{y})|}{|\alpha\boldsymbol{y}|}=0.$$

这就证明了所述论断.

特别，$Df(\boldsymbol{x})$在基向量 $\boldsymbol{e}_i$ 上取的值叫作 f 关于坐标变量 x_i 的'偏导数'，可写成

$$[Df(\boldsymbol{x})](e_i)=\lim_{\alpha\to 0}\frac{f(\boldsymbol{x}+\alpha\boldsymbol{e}_i)-f(\boldsymbol{x})}{\alpha}=\frac{\partial f}{\partial x_i}(\boldsymbol{x})$$

(其中 $i=1,\ \cdots,\ n$).

当然，$\frac{\partial f}{\partial x_i}(\boldsymbol{x})\in\mathscr{R}^m$ 为一函数向量，即下述列向量

$$\left(\frac{\partial f_1}{\partial x_i}(\boldsymbol{x}),\quad \frac{\partial f_2}{\partial x_i}(\boldsymbol{x}),\quad \cdots\quad,\ \frac{\partial f_m}{\partial x_i}(\boldsymbol{x})\ \right).$$

由上所述，只要假定 f 可微，那末它的偏导数都存在，并且就等于 $Df(\boldsymbol{x})$在各个基向量上取的值. 这里我们不加证明地指出，作为微分 $Df(\boldsymbol{x})$的线性映射，其矩阵形式为

$$\begin{pmatrix}\frac{\partial f_1}{\partial x_1}(\boldsymbol{x}), & \cdots, & \frac{\partial f_1}{\partial x_n}(\boldsymbol{x})\\ \cdots & \cdots & \cdots\\ \frac{\partial f_m}{\partial x_1}(\boldsymbol{x}), & \cdots, & \frac{\partial f_m}{\partial x_n}(\boldsymbol{x})\end{pmatrix},$$

这叫做 f 在点 x 的雅可比(C. G. J. Jacobi)矩阵. 特别，$m=1$ 时，记 $\Delta\boldsymbol{x}=(\Delta x_1,\ \cdots,\ \Delta x_n)$，则$[Df(\boldsymbol{x})]\Delta\boldsymbol{x}$ 称为 f 在 $\boldsymbol{x}$ 的全微分，即

$$[Df(\boldsymbol{x})]\Delta\boldsymbol{x}=\frac{\partial f(\boldsymbol{x})}{\partial x_1}\Delta x_1+\cdots+\frac{\partial f(\boldsymbol{x})}{\partial x_n}\Delta x_n.$$

5. 泰勒展开定理

泰勒展开定理的重要意义，在于表明：具有 $k+1$ 阶微分的函数 f，可以用 k 次多项式进行局部逼近.

先回顾一元函数的情形.

考虑映射 $f: \mathscr{R} \mapsto \mathscr{R}$. 我们知道，$f$ 在 x 点的导数就是微分 $Df(x)$在单位向量上的值. k 阶导数就是累次微分 $D^k f(x)$在单位向量上的值. 可用记号 $f'(x)$, $f''(x)$, …, $f^{(k)}(x)$来表示 f 的各阶导数.

泰勒定理 设 $G \subset \mathscr{R}$ 为开区间，$f: \mathscr{R} \mapsto \mathscr{R}$在 G上 $k+1$ 次可微. 则对于 x、$y \in G$，必有 x 和 y 之间的一点 ξ，使得

$$f(y)=f(x)+(y-x)f'(x)+\cdots+\frac{(y-x)^k}{k!}f^k(x)$$

$$+\frac{(y-x)^{k+1}}{(k+1)!}f^{k+1}(\xi).$$

证明 若 $x<y$，考虑区间$[x, y]$；否则，采用区间$[y, x]$. 引进如下的辅助函数

$$F(t)=f(y)-f(t)-(y-t)f'(t)-\cdots-\frac{(y-t)^k}{k!}f^k(t)$$

$$-\frac{(y-t)^{k+1}}{(k+1)!}\Big[f(y)-f(x)-(y-x)f'(x)$$

$$-\cdots\frac{(y-x)^k}{k!}f^k(x)\Big]\frac{(k+1)!}{(y-x)^{k+1}}.$$

它满足条件 $F(x)=F(y)=0$. 因此，应用微分中值定理的特例，即罗尔(M. Rolle)定理，可知有 $\xi \in (x, y)$，使得 $F'(\xi)=0$. 这就是说，

$$0=F'(\xi)=-\frac{f^{k+1}(\xi)(y-\xi)^k}{k!}$$

$$+\frac{(y-\xi)^k}{k!}\left[f(y)-f(x)-\cdots-\frac{f^k(x)}{k!}(y-x)^k\right]\cdot\frac{(k+1)!}{(y-x)^{k+1}}.$$

简化此式，便立即得到泰勒公式.

［特例］ 对于 $k=0$、$k=1$，分别有下列展式：

$$f(y)=f(x)+(y-x)f'(\xi),（其中\ x<\xi<y）.$$

$$f(y)=f(x)+(y-x)f'(x)+\frac{1}{2!}(y-x)^2f''(\xi),（其中x<\xi<y）.$$

如上的第一式便是微分中值公式.

6. 多元泰勒公式

设 $G\subset\mathscr{R}^n$ 为 n 维开区域，$f: G\mapsto\mathscr{R}$ 在 G 上 $k+1$ 次可微. 又设 $\boldsymbol{x}\in G$，$\boldsymbol{y}\in G$，且联结点 $\boldsymbol{x}$ 和点 $\boldsymbol{y}$ 的直线段位于G中. 在从 $\boldsymbol{x}=(x_1,\cdots,x_n)$到 $\boldsymbol{y}=(y_1,\cdots,y_n)$这个线段上我们得到一个变元 t 的函数 g，即

$$g(t)=f(x_1+t(y_1-x_1),\cdots,x_n+t(y_n-x_n)).$$

于是，在点 $\boldsymbol{x}+t(\boldsymbol{y}-\boldsymbol{x})$上计算偏导数，可得

$$g'(t)=\sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(y_i-x_i),$$

$$g''(t)=\sum_{i,j=1}^{n}\frac{\partial^2 f}{\partial x_i\partial x_j}(y_i-x_i)(y_j-x_j),$$

……………………………………

$$g^r(t)=\sum_{i_1,\cdots,i_r=1}^{n}\frac{\partial^r f}{\partial x_{i_1}\cdots\partial x_{i_r}}(y_{i_1}-x_{i_1})\cdots(y_{i_r}-x_{i_r}),$$

等等. 由一元函数的泰勒定理，

$$g(1)=g(0)+g'(0)+\cdots+\frac{1}{k!}g^k(0)+\frac{1}{(k+1)!}g^{k+1}(\xi),$$

其中 $0<\xi<1$. 代入，即得

$$f(\boldsymbol{y})=f(\boldsymbol{x})+\sum_{i=1}^{n}\frac{\partial f(\boldsymbol{x})}{\partial x_i}(y_i-x_i)$$

$$+\frac{1}{2!}\sum_{i,j=1}^{n}\frac{\partial^2 f(\boldsymbol{x})}{\partial x_i\partial x_j}(y_i-x_i)(y_j-x_j)+\cdots$$

$$+\frac{1}{k!}\sum_{i_1,\cdots,i_k=1}^{n}\frac{\partial^k f(\boldsymbol{x})}{\partial x_{i_1}\cdots\partial x_{i_k}}(y_{i_1}-x_{i_1})(y_{i_2}-x_{i_2})\cdots(y_{i_k}-x_{i_k})$$

$$+\frac{1}{(k+1)!}\sum_{i_1,\cdots,i_{k+1}=1}^{n}\frac{\partial^{k+1} f(\xi)}{\partial x_{i_1}\cdots\partial x_{i_{k+1}}}(y_{i_1}-x_{i_1})\cdots(y_{i_{k+1}}-x_{i_{k+1}}).$$

这就是 n 元实值函数的泰勒公式，它表明在固定点 $\boldsymbol{x}\in G$ 附近的函数值 $f(\boldsymbol{y})=f(y_1, \cdots, y_n)$ 可用多元多项式加以逼近．利用 $k=2$ 的特例，还可以讨论多元函数在 $\boldsymbol{x}$ 点取到相对极大、极小值的必要条件与充分条件．分析方法和一元函数情形相似，有兴趣的读者可以自行完成这种讨论．

六、黎曼积分及其推广

精通柯西的古典积分概念之后，人们自然会问：除了连续函数外，什么样的函数还可以进行积分？换句话说，对什么样的不连续函数，仍可引进定积分概念？这是德国数学家黎曼(G. F. B. Riemann)最先研究的问题，他的研究结果便引出‘黎曼积分’概念．

下面我们要讨论多元函数的黎曼积分，其实它并不比一元函数情形增加多少复杂性．为此，首先必须略述高维闭区间与开集等概念．

1．闭区间与开集

记 $\boldsymbol{x}=(x_1, \cdots, x_n)\in\mathscr{R}^n$．所谓 $\mathscr{R}^n$ 中的一个 n 维

闭区间，就是指点集

$$I=\{x\mid a_i\leqslant x_i\leqslant b_i,\ i=1,\ \cdots,\ n\},$$

其中 $a_i<b_i$. 自然，I 的体积 $\nu(I)$应由下式给出

$$v(I)=\prod_{i=1}^{n}(b_i-a_i)=(b_1-a_1)\cdots(b_n-a_n).$$

如果 $G\subset\mathscr{R}^n$ 是一个 n 维开集(它是 $\mathscr{R}$ 中开区间概念的拓广)，则有不相重叠的闭区间序列 $I_1,\ I_2,\ I_3,\ \cdots$，使得

$$G=\bigcup_{j=1}^{\infty}I_j=I_1\cup I_2\cup I_3\cup\cdots.$$

这里不证明这一事实了. 读者不妨将上述表示法的可能性视为开集的一个特征. 于是 G 的"体积"即定义为

$$v(G)=\sum_{j=1}^{\infty}v(I_j).$$

$\mathscr{R}^n$ 中有些点集是没有体积的，这时就规定其体积为零. 设 $S\subset\mathscr{R}^n$，如果对任给 $\varepsilon>0$，都存在有限个闭区间I_1、$\cdots$、I_m，使得

$$S\subset\bigcup_{j=1}^{m}I_j,\quad\sum_{j=1}^{m}v(I_j)<\varepsilon,$$

则称 S 为"零体积点集"，即 $\nu(S)=0$.

2. 有界函数的黎曼积分

设 I 为 n 维闭区间，$f\colon I\mapsto\mathscr{R}$ 为一有界函数. 所谓 I 的一个区间划分 $\pi=\{I_1,\ \cdots,\ I_s\}$，是指有限个不相重叠的闭区间集合，其并集为 I. 又若 $S\subset\mathscr{R}^n$，则点集 S 的直径是 $d(S)=\sup\{|x-y|\mid x\in S,\ y\in S\}$. 于是，划分 π 的模数定义为

$$\|\pi\|=\sup\{d(I_j)\mid j=1,\ \cdots,\ s\}.$$

对以下所论，读者当不难发现，其思想方法与柯西积分概念表述法有许多类似处，甚至关于相应命题的证明也十分相似.

如果对划分 $\pi=\{I_1, \cdots, I_s\}$ 中的区间 I_j 又进行了划分，则得到加细划分 $\pi'=\{J_1, \cdots, J_m\}$（这里 $s<m$）. 此时，称 π' 为 π 的'细分'，记作 $\pi'>\pi$，或 $\pi<\pi'$.

对 I 的每一划分 $\pi=\{I_1, \cdots, I_s\}$，引进黎曼和

$$U(f, \pi)=\sum_{j=1}^{s}\sup\{f(x)\mid x\in I_j\}\cdot v(I_j),$$

$$L(f, \pi)=\sum_{j=1}^{s}\inf\{f(x)\mid x\in I_j\}\cdot v(I_j).$$

显然 $L(f, \pi)\leqslant U(f, \pi)$. 由于在 π 的细分 π' 中，在那些更小的区间上 inf 可能增大，而 sup 可能减小，故知

$$\pi<\pi'\Rightarrow L(f,\pi)\leqslant L(f,\pi')\leqslant U(f,\pi')\leqslant U(f,\pi).$$

由此还可得一重要推论：若 π 与 π^* 是 I 的任意两个划分，则必 $L(f, \pi^*)\leqslant U(f, \pi)$. 事实上，只须取一个共同的细分 $\pi''>\pi^*$，$\pi''>\pi$，就得出

$$L(f, \pi^*)\leqslant L(f, \pi'')\leqslant U(f, \pi'')\leqslant U(f, \pi).$$

由于划分的细分过程总是使 L 上升而 U 下降，故对取所有划分 $\{\pi\}$ 而言，按确界存在定理，可引进如下的高积分与低积分概念

$$\overline{\int} f=\inf\{U(f, \pi)\mid\pi\}, \qquad \underline{\int} f=\sup\{L(f, \pi)\mid\pi\}.$$

综上所述，我们已证明下述辅助命题：

引理 对于每一个有界函数 $f: I\mapsto\mathscr{R}$，恒有

$$\overline{\int} f \geqslant \underline{\int} f,$$

而且,等号成立的必要与充分条件是:对于任给 $\varepsilon>0$, 总有 I 的划分 π,使得 $U(f,\pi)-L(f,\pi)<\varepsilon$.

特别, 如果$\overline{\int} f=\underline{\int} f$,则称 f 在 I 上为'黎曼可积',可记为$\int f$或$\int_I f(x)dx$.

注意, 一元函数的一致连续性概念可以推广到多元函数,而且不难类似地证明:"凡是在闭区间 I 上连续的 n 元函数,一定是一致连续的."由此立即可证重要命题:"若 f: $I\mapsto\mathscr{R}$ 连续,则 f 为可积".

证明 任给 $\varepsilon>0$.由一致连续性可知:有 $\delta>0$, 当 x、$y\in I$,且 $|x-y|<\delta$ 时,恒有

$$|f(y)-f(x)|<\frac{\varepsilon}{v(I)}.$$

设 $\pi=\{I_1,\cdots,I_s\}$是 I 的一个划分,模数 $\|\pi\|<\delta$.则

$$\begin{aligned}U(f,\pi)-L(f,\pi)&=\sum_{j=1}^{s}\sup\{f(x)\mid x\in I_j\}v(I_j)\\&\quad-\sum_{j=1}^{s}\inf\{f(x)\mid x\in I_j\}v(I_j)\\&=\sum_{j=1}^{s}[\sup\{f(x)|x\in I_j\}-\inf\{f(x)|x\in I_j\}]v(I_j)\\&\leqslant\frac{\varepsilon}{v(I)}\sum_{j=1}^{s}v(I_j)=\varepsilon.\end{aligned}$$

故由引理,便证得 f 为黎曼可积.

3. 黎曼可积的充要条件

勒贝格于1904年在《积分与原函数研究》的著名论文

中证明了：有界函数黎曼可积的必要与充分条件是，不连续点构成一个零体积点集．这就彻底解决了黎曼可积性问题．下面就闭区间 $I\in\mathscr{R}^n$ 上的有界函数 f 来论述这一问题．先定义 f 在一点 $x\in I$ 的'振幅'．对于每一个$\delta>0$，记

$$W_\delta(x)=\sup\{f(y)\,|\,|y-x|\leqslant\delta\}-\inf\{f(y)\,|\,|y-x|\leqslant\delta\}.$$

于是，f 在 x 的'振幅'就是下列极限值

$$W(x)=\lim_{\delta\to 0}W_\delta(x).$$

显然，只有当 f 在 x 点连续时$W(x)=0$．一般情形$W(x)\geqslant 0$．

现在，利用我们在前面得到的引理，就可以来建立关于黎曼可积性的特征定理："设 $f:I\mapsto\mathscr{R}$ 为有界函数，那末，当且仅当对于每一个 $k>0$，全体使 f 的振幅大于或等于 k 的点成一零体积集合时，f 为黎曼可积．"

证明 假设定理中的条件成立．任意给定$\varepsilon>0$．考虑 I 中的不连续点的点集（当 f 连续时，此点集为空集）：

$$S_\varepsilon=\{x\,|\,W(x)\geqslant\varepsilon\}.$$

由于 $\nu(S_\varepsilon)=0$，故 S_ε 可为 I 中的不相重叠的闭区间所覆盖，而这些区间的体积之和小于$\frac{\varepsilon}{2M}$，此处M为$|f(x)|$在 I 上的上确界．这些区间可以微微地延拓成开区间J_1，…，J_r，而使其体积之和仍小于$\frac{\varepsilon}{2M}$．于是，余集$I-\bigcup_{i=1}^{r}J_i$为一闭集，f 在此闭集上各点的振幅均小于 ε．类似于

证明一致连续性那样，可以找到关于该余集的一个划分 $\{I_1, I_2, \cdots, I_s\}$，使得对于 x、$y \in I_j$（其中 $j=1, \cdots, r$），恒有 $|f(x)-f(y)|<\varepsilon$.

再任取 I 的划分 $\pi=\{I_1, I_2, \cdots, I_s, I_{s+1}, \cdots, I_t\}$. 于是，根据黎曼和的定义，易见

$$U(f, \pi)-L(f, \pi) \leqslant \varepsilon \cdot v(I)+2M \cdot(\varepsilon / 2M)=(1+v(I)) \varepsilon.$$

因可使 $\varepsilon \rightarrow 0+$，故由引理得知 f 为黎曼可积.

反之，假如定理条件不满足. 则存在一个 $k>0$，使点集 $S_k=\{x \mid W(x) \geqslant k\}$ 的体积 $\nu(S_k)>0$. 因此存在 $\delta>0$，只要当 $I_1, \cdots, I_r$ 覆盖 S_k 时，总有

$$\sum_{j=1}^{r} v(I_j)>\delta.$$

今设 $\pi=\{I_1, \cdots, I_m\}$ 是 I 的任意划分，其中 $I_1, \cdots, I_r$ 是内点中含有 S_k 之点的那些区间. 故如上形式的不等式总成立. 因此

$$U(f, \pi)-L(f, \pi) \geqslant k \cdot \sum_{j=1}^{r} v(I_j)>k \delta.$$

这表明必然是 $\overline{\int} f>\underline{\int} f$. 故 f 不可积. 定理得证.

［特例］ 在一维区间 I 内，任意有限多个点的体积为零；在二维区间 I 内，任意有限多条直线的体积为零；余仿此. 因此，凡有界的一元函数，不论有多少个（有限多个）不连续点，它总是黎曼可积的. 凡有界的二元函数，即使在 $I \in \mathscr{R}^2$ 内的任意有限多条直线段上不连续，也不损害它的黎曼可积性. 其余仿此类推.

4. 开集上的黎曼积分

设 $G\subset\mathscr{R}^n$ 为 n 维开集，$f:G\mapsto\mathscr{R}$ 为非负函数，且在 G 中的每一个闭区间上都可积．这就导致 f 必定在 G 中的每个闭区间上是有界的，但 f 在整个 G 上却可能是无界的．

设 $I_1, I_2, \cdots$ 为不相重叠的闭区间，其并集为 G，且使 G 中的每个有界闭集都包含于其有限个的并集中．那末非负函数 f 在 G 上的积分便定义为

$$\int_G f=\sum_{k=1}^{\infty}\int_{I_k} f.$$

这个积分值是非负的，也可能取值 $+\infty$（正无限）．

对于一般的函数 f，可分解为两个非负函数之差．令 $f^+(x)=\max\{f(x),0\}$，$f^-(x)=-\min\{f(x),0\}$．于是 $f(x)=f^+(x)-f^-(x)$，且 f^+ 和 f^- 都是非负函数．这里不加证明地指出："设 f 是闭区间 I 上的函数，那末，当且仅当 f^+ 和 f^- 都可积时，f 为可积；此时，有 $\int f=\int f^+-\int f^-$．"

根据上述命题，即可给出一般函数的黎曼积分定义："设 $f: G\mapsto\mathscr{R}$，此处 $G\subset\mathscr{R}^n$ 为一开集．如果 f 在每一个闭区间 $I\subset G$ 上都可积，且积分 $\int_G f^+$ 与 $\int_G f^-$ 中至少有一个取有限值，那末就说 f 在 G 上可积．" 只要 f 在 G 上可积，就可写成

$$\int_G f=\int_G f^+-\int_G f^-.$$

当然这也是一个定义．据此，若 $\int_G f^+=\infty$，而 $\int_G f^-<\infty$，则定义 $\int_G f=\infty$；若 $\int_G f^+<\infty$ 而 $\int_G f^-=\infty$，则定义 $\int_G f=-\infty$．

顺便指出，在现代的积分论著作中，常把 $+\infty$ 与 $-\infty$ 视为广义实数；但这只能看成是一种方便的说法，由于它们不在数域中，故不能任意作运算．至于建立了'非标准数域'之后，那当然是另外一回事了．

5．黎曼积分的一个推广

作为黎曼积分的一个简单而自然的拓广，就是'斯蒂尔吉斯(Th. J. Stieltjes)积分'．这里，只就一元函数的情形，略述其大意．

记 $I=[a,\ b]$，设 $f:\ I\mapsto\mathscr{R}$ 和 $g:\ I\mapsto\mathscr{R}$ 为给定的二个函数．作划分 $\Delta:\ a=x_0<x_1<\cdots<x_n=b$，其模数为 $\|\Delta\|$．任意选取 $\xi_i\in I_i=[x_i,x_{i+1}]$（其中 $i=0,1,\cdots,n-1$）．如果存在下述极限

$$\lim_{\|\Delta\|\to 0}\sum_{i=0}^{n-1}f(\xi_i)[g(x_{i+1})-g(x_i)],$$

并且，其极限值与划分方式无关，还和点 ξ_i 的选法无关，那末这个极限值就叫作函数 $f(x)$ 关于 $g(x)$ 从 a 到 b 的"斯蒂尔吉斯积分"，记作 $\int_a^b f(x)dg(x)$．

往下我们假定 g 是 I 上的有界的单调非降函数，从而 $g(x_{i+1})-g(x_i)\geqslant 0$，而且

$$\sum_{i=0}^{n-1}[g(x_{i+1})-g(x_i)]=g(b)-g(a).$$

这样一来，几乎完全可以照套前面关于论述黎曼积分存在性的思想方法，去得出某些相类似的命题．

对于 $I=[a,\ b]$ 上的有界函数，仿黎曼和的形式，记

$$U(f,\ \Delta)=\sum_{k=0}^{n-1}\sup\{f(x)\,|\,x\in I_k\}[g(x_{k+1})-g(x_k)],$$

$$L(f,\ \Delta)=\sum_{k=0}^{n-1}\inf\{f(x)\mid x\in I_k\}[g(x_{k+1})-g(x_k)].$$

如果 Δ' 是 Δ 的细分，即 $\Delta'>\Delta$，则同理可得

$$L(f,\ \Delta)\leqslant L(f,\ \Delta')\leqslant U(f,\ \Delta')\leqslant U(f,\ \Delta).$$

考虑取遍所有划分时，应用确界存在定理同样地可以定义高、低积分

$$\overline{\int} f\,dg=\inf\{U(f,\ \Delta)\mid\Delta\},\quad \underline{\int} f\,dg=\sup\{L(f,\ \Delta)\mid\Delta\}.$$

显然 $\overline{\int} f\,dg\geqslant\underline{\int} f\,dg$. 特别，当等号成立时，便得到斯蒂尔吉斯积分

$$\overline{\int} f\,dg=\underline{\int} f\,dg=\int_a^b f(x)\,dg(x).$$

因此，类似于黎曼积分的推理，我们也有如下的一个基本命题.

引理 斯蒂尔吉斯积分存在的必要与充分条件是：对于任给 $\varepsilon>0$，总有 I 的一个划分 Δ 使得 $U(f,\ \Delta)-L(f,\ \Delta)<\varepsilon$. 也就是说，

$$\lim_{\|\Delta\|\to 0}[U(f,\ \Delta)-L(f,\ \Delta)]=0.$$

应用这引理，即可证得重要定理："若 $f(x)\in\mathscr{C}(I)$，则积分 $\int_a^b f(x)\,dg(x)$ 必存在."

证明 由一致连续性，对任给 $\varepsilon>0$，可知总存在 $\delta>0$，使当 $\|\Delta\|<\delta$ 时，恒有

$$\sup\{f(x)\mid x\in I_k\}-\inf\{f(x)\mid x\in I_k\}<\varepsilon,$$

其中 $k=0,1,\ \cdots,\ n-1$. 于是

$$0\leqslant U(f,\ \Delta)-L(f,\ \Delta)\leqslant \varepsilon[g(b)-g(a)].$$

这表明 $\|\Delta\|\to 0\Rightarrow[U(f,\ \Delta)-L(f,\ \Delta)]\to 0$. 故定理为真.

6. 有界变差函数的积分

如果函数 $g:I=[a,b]\mapsto\mathscr{R}$ 能表示成两个有界的单调非降函数之差，如

$$g(x)=\phi(x)-\psi(x),\qquad (x\in I)$$

此处 $\phi(x)$ 与 $\psi(x)$ 均为非降函数，那末就称 g 为**有界变差函数**. 由此定义立即推知：任何两个有界变差函数的和、差及乘积也都是有界变差函数.

利用上面已经证过的定理，可推得下列二论断(也可作为定理看待)：

(i) 若 $g(x)$ 为 I 上的有界变差函数，而 $f\in\mathscr{C}(I)$，则 f 关于 g 在 I 上的斯蒂尔吉斯积分存在.

事实上，如果 $g=\phi-\psi$，其中 ϕ 与 ψ 为非降函数，则

$$\int_a^b f(x)\,dg(x)=\int_a^b f(x)\,d\phi(x)-\int_a^b f(x)d\psi(x).$$

(ii) 若 $f(x)$ 为 I 上的有界变差函数，而 $g\in\mathscr{C}(I)$，则 f 关于 g 的斯蒂尔吉斯积分存在，而且有如下的分部积分公式

$$\int_a^b f(x)dg(x)=f(b)g(b)-f(a)g(a)-\int_a^b g(x)df(x).$$

注意，在上述公式中，f 与 g 的地位是对称的. 在它们中，只要有一个是连续的，一个是有界变差的，那末 f 关于 g 以及 g 关于 f 的一对斯蒂尔吉斯积分都必同时存在.

现在来证明论断(ii)：

任取 I 的划分 $\Delta=\{I_0, I_1, \cdots, I_{n-1}\}$，对和式

$$\sigma(f, \Delta)=\sum_{k=0}^{n-1} f(\xi_k)[g(x_{k+1})-g(x_k)], \qquad (\xi_k\in I_k)$$

利用阿贝尔的分部求和法，可将它改写成

$$\begin{aligned}\sigma(f, \Delta)=-\sum_{k=1}^{n-1} g(x_k)[f(\xi_k)-f(\xi_{k-1})]-g(a)[f(\xi_0)\\ -f(a)]-g(b)[f(b)-f(\xi_{n-1})]\\ +f(b)g(b)-f(a)g(a).\end{aligned}$$

注意 f 是有差变差函数，故根据论断(i)，g 关于 f 的积分是存在的，从而上式右端含有 $\xi_0, \xi_1, \cdots, \xi_{n-1}$ 各点的诸项的和式(共有 $n+1$ 项)于 $\|\Delta\|\to 0$ 时以 $\int_a^b g\,df$ 为极限. 这说明 $\sigma(f, \Delta)$的极限是存在的，通过极限手续便得出

$$\int_a^b f\,dg=-\int_a^b g\,df+f(b)g(b)-f(a)g(a).$$

故知论断(ii)成立.

7. 广义积分

有界变差函数 $g:I\mapsto\mathscr{R}$ 如果在 $I=[a, b]$上符合条件

$$g(a)=0,$$

$$g(x)=\frac{1}{2}[g(x+)+g(x-)], \qquad (a<x<b)$$

就称 g 为正规化有界变差函数.

考虑无限区间 $J=[a,\infty)$. 设$f\in\mathscr{C}(J)$，又设对于每个 $M>0$,$g(x)$是$[a,M]$上的正规化有界变差函数. 于是 f 关于 g 在$[a, \infty)$上的‘广义积分’即定义为

$$\int_a^\infty f(x)\,dg(x)=\lim_{M\to\infty}\int_a^M f(x)dg(x).$$

当上述极限存在时，称积分为收敛，否则称为发散．同样可规定$[-\infty, a)$上广义积分的概念．又，$(-\infty, \infty)$上的广义积分可由下式定义

$$\int_{-\infty}^{\infty} f(x)\, dg(x) = \int_{-\infty}^{0} f(x)\, dg(x) + \int_{0}^{\infty} f(x) dg(x),$$

这里假定$f \in \mathscr{C}(\mathscr{R})$，又$g(x)$在任何有限区间$[\alpha, \beta] \subset \mathscr{R}$上都是有界变差函数．

最后，值得补充指出，积分是一种兼具有计算与表现双重职能的数学工具．例如在现代泛函分析中，积分常被用来表现若干类重要的线性泛函，或者作为解决微分方程等问题所需要的线性变换工具．在应用上特别重要的线性映射，如傅里叶(J. B. J. Fourier)变换、拉普拉斯(P. S. M. de Laplace)变换等，都是借助于特种形式的广义积分来表现的．这些积分变换所定义的函数类，已被拓广成‘广义函数’类．正如1984年出版的《美国数学的现在和未来》(Renewing U. S. Mathematics，周仲良译，复旦大学出版社 1986 年版)一书中指出的："积分变换的另一方向上的推广是由 1916—1917 年左右由丰克(P. Funk)和拉顿(Raden)做出来的．那是将函数$f(x)$的变换定义为沿一条曲线(点集)上的积分，而原来的f可从它的拉顿变换中重新造出来．经过60来年，物理学家柯麦克(Cormack)在题为《函数的线积分表示法》论文中阐明了如何根据 X 射线的观测再造出图象来．这一思想在实践中的发展导致了计算机辅助层面摄影的出现，并获得1979年诺贝尔医学奖．"

七、函 数 级 数

“函数级数”就是由一串(通常是无穷多个)函数相加构成的级数．这种级数既是表现函数的工具，又是用来进行计算的工具．在微积分学中，主要考虑的问题是：在什么样的条件下，函数级数可以逐项积分和逐项求导？

具体地说，如果假设各个 $u_k(x)$ 都是 $I=[a, b]$ 上的连续函数(其中 $k=1, 2, 3, \cdots$)，且有如下的级数和

$$\sum_{k=1}^{\infty} u_k(x) = u_1(x) + u_2(x) + u_3(x) + \cdots = S(x),$$

这里的 $S(x)$ 叫做级数的“和函数”，那末试问在什么条件下能保证可以进行如下所示的逐项积分：

$$\int_a^b S(x)\,dx \equiv \int_a^b \left(\sum_{k=1}^{\infty} u_k(x)\right)dx = \sum_{k=1}^{\infty}\int_a^b u_k(x)dx \ ?$$

又假如各个 $u_k(x)$ 都是 (a, b) 上的可微函数，则可以问在什么条件下能进行逐项求导：

$$S'(x) \equiv \frac{d}{dx}\sum_{k=1}^{\infty} u_k(x) = \sum_{k=1}^{\infty}\frac{d}{dx}u_k(x) \equiv \sum_{k=1}^{\infty} u_k'(x)?$$

上述两个问题，19世纪20年代初在柯西出版的《分析教程》名著中已有所论述，但由于当时柯西的头脑里还未形成“一致收敛性”概念，以致他曾依靠直观作出了一些错误的论断．

1．阿贝尔的反例

柯西的一个主要疏忽是，他误认为由连续函数项作

成的级数 $\sum_{k=1}^{\infty} u_k(x)$ 之和 $S(x)$ 也一定是连续函数. 另一点疏忽是,他误以为上述的函数级数总可以逐项积分,而忽视了"一致收敛性"的要求. 当年挪威的青年数学家阿贝尔 (N. H. Abel) 已经通晓傅里叶级数概念,他知道函数 $S(x)=\frac{x}{2}$ 在开区间 $(-\pi, \pi)$ 中的"傅里叶展开"是

$$\sin x-\frac{\sin 2x}{2}+\frac{\sin 3x}{3}-\frac{\sin 4x}{4}+\cdots.$$

因为这是一个以 2π 为周期的函数级数,所以,它表示了一个周期函数,从而,它在每个 2π 长的区间内的图形和 $(-\pi, \pi)$ 内的函数 $y=\frac{x}{2}$ 图形(直线段)是一致的. 因此,在整个区间 $(-\infty, \infty)$ 内函数级数所代表的和函数的图形,便是一串以 $\frac{1}{2}$ 为斜率的折线段.

这样,上述的连续函数项作成的级数在每个点 $x=(2k+1)\pi$ 处都是不连续的. 事实上,其左、右极限分别为 $S((2k+1)\pi-)=\frac{\pi}{2}$, $S((2k+1)\pi+)=-\frac{\pi}{2}$, 其中 $k\in\mathbf{Z}$.

阿贝尔曾在 1826 年的著名论文中举出了上述反例,并婉转地指出了柯西的疏忽. 他写道:"可惜柯西先生的结论未能排除例外.……"在阿贝尔的论文中,他还利用一致收敛的思想,正确地证明了"连续函数为项的一个一

致收敛级数的和，在收敛区域内部是连续的."可惜阿贝尔当时没有从中把一致收敛的性质抽象概括出来，形成普遍概念.

一致收敛的一般概念，是在 1842 年由外尔斯特拉斯提出的. 英国的第一流数学物理学家斯托克斯 (G. G. Stokes) 也在1848年发现这一概念. 差不多同一时间，德国的赛德尔(W. P. Seidel) 也认识到这一重要概念. 至于分析学奠基者柯西本人，最后也还是认识到，为了断言连续函数的级数的和一定连续，需要一致收敛性条件. 这与他出版名著《分析教程》时已相隔30年了.

2. 一致收敛性概念

给定某个区间 I 上的一个函数级数$\sum_{1}^{\infty} u_k(x)$ (其中 $x\in I$). 如果对每点 $x\in I$，都有

$$\sum_{k=1}^{\infty} u_k(x)=S(x), \quad \text{或} \quad \lim_{n\to\infty}\sum_{k=1}^{n} u_k(x)=S(x),$$

则称该函数级数逐点收敛到 $S(x)$. 用 ε-N 准则来说，便是：对每个 $x\in I$，任给 $\varepsilon>0$，总有一个正数 $N=N(\varepsilon, x)$，通常依赖于 ε 和 x，使得

$$n>N=N(\varepsilon,x) \quad \Rightarrow \quad \left|S(x)-\sum_{k=1}^{n} u_k(x)\right|<\varepsilon.$$

所谓一致收敛性，就是说：对任给 $\varepsilon>0$，总有一个不依赖于 x 的正数 $N=N(\varepsilon)$，它只与 ε 有关，使得对一切 $x\in I$ 都有

$$n>N=N(\varepsilon) \quad \Rightarrow \quad \left|S(x)-\sum_{k=1}^{n} u_k(x)\right|<\varepsilon.$$

因为如上刻画趋近速度的关系式 (不等式) 对全体

$x\in I$ 是一致地成立的，所以也就称它为一致收敛性．可以简记成 $\sum_{k=1}^{n}u_k(x)\xrightarrow{\text{1}}S(x)$，$x\in I$，$(n\to\infty)$．

一致收敛性的几何意义是很明显的．设想和函数 $y=S(x)$（其中 $x\in I$）是一条曲线（可以是不连续的），$S_n(x)=\sum_{k=1}^{n}u_k(x)$（其中 $n=1, 2, 3, \cdots$）是一系列曲线．两条曲线

$$y=S(x)-\varepsilon,\qquad y=S(x)+\varepsilon\qquad (x\in I)$$

构成坐标平面上的一条带状区域（或许有间断），可称它为 $y=S(x)$ 的"ε-走廊"．一致收敛性 $S_n(x)\xrightarrow{\text{1}}S(x)$ 的意思就是说：只要 $n>N=N(\varepsilon)$，则一切曲线 $y=S_n(x)$ 便全部进入 $y=S(x)$ 的"ε-走廊"内，也即

$$|S(x)-S_n(x)|<\varepsilon.\qquad (n>N,\ x\in I)$$

读者或许会提问：在逐点收敛性定义中已经有 $N(\varepsilon, x)$．为了找到一个适用于一切 x 的 $N(\varepsilon)$，可否取

$$N(\varepsilon)\ :=\ \sup\{N(\varepsilon,\ x)\,|\,x\in I\}?$$

当然，如果上面的上确界确实是个有限数，则如上选取的 $N(\varepsilon)$ 便符合一致收敛性要求．换句话说，当上述上确界对每个 $\varepsilon>0$ 都存在时，函数级数在 I 上必定是一致收敛的．但如果 $\sup\{N(\varepsilon,\ x)\}=+\infty$，则将是另一回事了．读者还可以进一步思考一下："非一致收敛性"应该怎样表述明白．

通常，把不等式 $|S(x)-S_n(x)|<\varepsilon$（其中 $x\in I$）称之为"部分和 $S_n(x)$ 对和函数的逼近偏差为 ε"；因此，$S_n(x)$ 对 $S(x)$ 的一致收敛性在函数逼近论中又叫作"一致逼近

性[▼]. 如果 $S(x)$ 也是连续函数,则

$$\max_{x\in I}|S(x)-S_n(x)|\leqslant\varepsilon.$$

这时 ε 便叫作 $S_n(x)$ 对于 $S(x)$ 的偏差(或逼近度). 显然,逼近度是个宏观概念,它是对整个区间 I 上的函数而言的.

3. 连续函数的级数连续性定理

如果某区间 I 上的连续函数的级数 $\sum_{k=1}^{\infty}u_k(x)$ 一致收敛到 $S(x)$,那末,和函数 $S(x)$ 也是连续函数.

证明 只须证明对每点 $x_0\in I$ 都有

$$x\to x_0\ (\text{其中 } x\in I)\quad\Rightarrow\quad|S(x)-S(x_0)|\to 0.$$

用 ε-δ 准则来表述,便是:对任给 $\varepsilon>0$,必有 $\delta>0$,使得

$$|x-x_0|<\delta\quad\Rightarrow\quad|S(x)-S(x_0)|<\varepsilon.$$

记 $S_n(x)=\sum_{k=1}^{n}u_k(x)$. 显而易见

$$\begin{aligned}|S(x)-S(x_0)|&=|S(x)-S_n(x)+S_n(x)-S_n(x_0)+S_n(x_0)-S(x_0)|\\&\leqslant|S(x)-S_n(x)|+|S_n(x)-S_n(x_0)|+|S_n(x_0)\\&\quad-S(x_0)|.\end{aligned}$$

按一致收敛性. 可选取 $N=N(\varepsilon)>0$,使得

$$n\geqslant N\quad\Rightarrow\quad|S(x)-S_n(x)|<\frac{1}{3}\varepsilon,$$

$$\Rightarrow\quad|S(x_0)-S_n(x_0)|<\frac{1}{3}\varepsilon.$$

固定选取 $n=N$,则 $S_N(x)$ 为连续函数,故有 $\delta>0$,使得

$$|x-x_0|<\delta\quad\Rightarrow\quad|S_n(x)-S_n(x_0)|<\frac{1}{3}\varepsilon.$$

从而

$$|S(x)-S(x_0)|<\frac{1}{3}\varepsilon+\frac{1}{3}\varepsilon+\frac{1}{3}\varepsilon=\varepsilon.$$

由此证明了函数 $S(x)$的连续性.

4. 逐项积分定理

这个定理就是断言:"凡是在区间 $I=[a, b]$ 上一致收敛的连续函数的级数 $\sum_{k=1}^{\infty}u_k(x)=S(x)$,一定是可以逐项积分的."就是说

$$\int_a^b S(x)\,dx=\sum_{k=1}^{\infty}\int_a^b u_k(x)dx.$$

证明 对任给 $\varepsilon>0$,按一致收敛性,可得 $N=N(\varepsilon)>0$,使得当 $n>N$ 时,保证有

$$|r_n(x)|\equiv\left|S(x)-\sum_{k=1}^{n}u_k(x)\right|<\frac{\varepsilon}{b-a}. \qquad (x\in I)$$

因此

$$\left|\int_a^b S(x)dx-\sum_{k=1}^{n}\int_a^b u_k(x)dx\right|=\left|\int_a^b\left(S(x)-\sum_{k=1}^{n}u_k(x)\right)dx\right|$$

$$=\left|\int_a^b r_n(x)dx\right|\leqslant\int_a^b|r_n(x)|dx<\varepsilon. \qquad (n>N)$$

最后使 $\varepsilon\to0+$(相应地 $n\to\infty$),即得

$$\int_a^b S(x)dx=\lim_{n\to\infty}\sum_{k=1}^{n}\int_a^b u_k(x)dx.$$

这样便证明了逐项积分定理.

5. 逐项求导定理

这个定理要求逐项求导后的级数在一个区间内是一致收敛的,因而很容易地从逐项积分定理把它推导出来.求导定理是说:"如果在区间(a, b)内 $S(x)=\sum_{k=1}^{\infty}u_k(x)$,

而且$\sum\limits_{k=1}^{\infty} u'_k(x)$在$(a, b)$内是一致收敛的连续函数的级数,那末就有$S'(x)=\sum\limits_{k=1}^{\infty} u'_k(x)$."

证明 设$a<x_0<x<b$.显然可写

$$S(x)-S(x_0)=\sum_{k=1}^{\infty}[u_k(x)-u_k(x_0)].$$

由积分学基本定理,可知上式又可写成

$$S(x)-S(x_0)=\sum_{k=1}^{\infty}\int_{x_0}^{x} u'_k(t)dt.$$

又因在(a, b)内$\sum\limits_{k=1}^{\infty} u'_k(t)$是一致收敛的,故由逐项积分定理可知上式也可表成

$$S(x)-S(x_0)=\int_{x_0}^{x}\sum_{k=1}^{\infty} u'_k(t)dt.$$

由此可见,上式两端都是x的可微函数,因此上式等价于

$$\int_{x_0}^{x} S'(t)dt=\int_{x_0}^{x}\sum_{k=1}^{\infty} u'_k(t)dt.$$

最后,对两端求导,就得到要证明的结论.

在一般微积分教科书中,可以找到大量例子,以说明逐项积分与逐项求导定理的具体应用.

6. 级数连续性的充要条件

注意前面我们在证明"级数连续性定理"时,实际上只用到了级数在每点邻域内的一致收敛性.这就启发人们去寻找函数级数连续性的充要条件.这一研究课题是由意大利的分析学家阿尔齐拉(C. Arzela)在19世纪末完成的.他首先引进了"次一致收敛性"(Sub-uniform

convergence)概念，即下述定义：

对于任给 $\varepsilon>0$ 及正整数 m，如果区间 $I=[a, b]$ 总可以由一批小的开区间 $I_{1(\varepsilon)}, I_{2(\varepsilon)}, \cdots, I_{p(\varepsilon)}$ 所盖满；对于这些小区间，又相应地存在一批全都大于 m 的整数 $N_{1(\varepsilon)}, N_{2(\varepsilon)}\cdots, N_{p(\varepsilon)}$，使得

$$x\in I_{j(\varepsilon)},\ n=N_{j(\varepsilon)} \quad\Rightarrow\quad |r_n(x)|=\left|\sum_{n+1}^{\infty}u_k(x)\right|<\varepsilon,$$

此处 $j=1, \cdots, p$. 这样，就说级数 $\sum_1^{\infty}u_k(x)$ 在$[a, b]$上为**次一致收敛**.

有了这个概念，就可以把全区间上的一致收敛性条件减弱成为各个局部小区间上的收敛性要求.于是，经过精心分析，阿尔齐拉便证明了下述定理：

"设 $\sum_{k=1}^{\infty}u_k(x)=S(x)$，其中各个 $u_k(x)$ 都是区间$[a, b]$上的连续函数. 那末，和函数 $S(x)$在$[a, b]$上连续的必要与充分条件是：该函数级数在$[a, b]$上为次一致收敛."

为节约篇幅，这里就不给出它的详细证明了.读者欲知其详，建议查阅实变函数论专书. 最后我们指出，逐项积分定理的重要拓广是由勒贝格(H. L. Lebesgue) 在本世纪之初作出的. 他建立的一般逐项积分定理(又名"有界收敛定理")，是针对可测函数序列作成的级数而言的，那也是近、现代实变函数论中的主要论题之一.甚至在勒贝格1904年出版的《关于积分法与原函数研究的讲义》一书里，曾把他的上述定理看成为他的全部理论的基石，其

重要意义可想而知.

八、非标准分析大意

非标准分析是1961 年美国数理逻辑学家A·鲁宾逊(A. Robinson)开始创立的一门新的分析学. 他在 1966 年出版的《非标准分析》一书，可以看成是这一崭新数学分支的奠基性著作.此书于1980年已由申又枨、王世强等完成中译本,由科学出版社出版.

20多年来,非标准分析已有许多发展.但这里只能用很短的篇幅,就其基本思想方法,作一简略介绍；相信读者当能由此略见其真貌.

1. 一点历史回顾

由分析学发展史我们知道,牛顿、莱布尼茨创立微积分时曾借用“无限小”概念. 但是他俩的思维背景和思路是很不相同的. 牛顿的创造性思想背景是力学. 所以他始终坚持了变量的连续性特征. 而莱布尼茨在哲学上有他一套“单子论”宇宙观.所以在建立微积分时,他曾有意识地发展了无限小概念，企图把微积分搞成真正的无限小分析,而且在心目中把无限小看成为思维中的“单子”. 这样的思想方法，就促使他引进了特别便于应用的记号 dy、dx 和记法$\frac{dy}{dx}$等. 当然,这些记号的精确含义,还是等到极限论出来后,才变得完全明确的.

正因为实数是满足阿基米德(Archimedes) 公理的，即"设 $0<a<b$，总存在自然数 n，使得 $na>b$."但无限小的任何整数倍还是无限小，不可能大于任何正实数，所以无限小并非通常的实数. 换言之，实数域 **R**（以前记为 $\mathscr{R}$）中并不存在什么"无限小". 因此，一切利用无限小作为推导数学命题的方式，都是没有根据的，自然是不能令人信服的，又何况形式上用无限小进行推理时还会产生自相矛盾的"悖论".

所以，自从后来有了极限论作为微积分理论基础后，实际上无限小就被废黜了；以致如有人再想用无限小进行推理时，甚至被认为开"历史倒车". 可是，又不可否认分析学中确实有大批命题借助于直观上的无限小，能一眼看到底. 至少，应用无限小概念可以较快地预见到许多客观上能被逻辑地证实的结论. 因此，还是有少数独具慧眼的数学家（包括晚年时期的希尔伯特(D. Hilbert)），都曾设想过如何应用现代数学逻辑成就，使当年的"莱布尼茨之梦"在新的合理基础上重新复活. 这一历史任务，终于通过鲁宾逊及其后继者的努力，彻底实现了.

建立非标准分析的关键问题，就是如何把标准实数域 **R** 扩张成为非标准数域***R**，使得那些无限小与无限大（所谓"非标准实数"）也成为***R** 中的合法成元，从而可以进行各种运算. 固然对***R** 来说，不能再要求它保持阿基米德公理了. 但为了建立分析学，却必须要求它是个有序域. 所以***R** 应是 **R** 的一个有序扩张，而 **R**⊂***R** 是个真子域. 鲁宾逊的重要贡献之一，就是他成功地运用模

型论方法，示明了构造$^*R$模型的一般途径．但是，他使用的逻辑工具十分繁重(例如，在他的奠基性著作中，差不多用了四十页篇幅，来铺设逻辑框架)．这就正如70年代法国数学家指出的，由于多数数学家对过份沉重的逻辑框架不感兴趣，以致就不乐愿问津非标准分析了．非标准分析至今在国际数学界尚未真正普及，恐怕这也是一个主要原因．

2．非标准分析的特点

就实质而言，非标准分析与标准分析是等价的，即前者所能证明的命题也必能为后者所证明，反之亦然．这一事实也常常成为不乐意学习非标准分析者的借口．事实上，非标准分析有三大特点是很值得重视的：一是非标准数域$^*R$比之标准实数域R有更丰富的内涵，这就为应用上提供更广阔和更有表现力的空间模式．二是在分析论证与计算方式上，它比标准分析直捷了当些．这是因为标准分析采用的极限论方法是分两步走的方法，即先是逐步近似，然后再完成极限过程；而非标准分析直接使无限小、无限大等广义实数参与运算，故本质上是一步走的方法．因此，如能精通此种方法，则对发现新结果将会提供方便．三是非标准区间的无限小分划具有貌似离散化的性质，这就成为离散数学(如组合数学)技巧能转移到连续数学领域的一个桥梁．而这一过渡桥梁，却是标准分析所不能提供的．正是这些特点，可以加强人们对哥德尔预言的信念：即非标准分析必将成为“未来的分析学”．

3. 关于*R 的高桥模型

构造非标准数域 *R 模型的方法很多.事实上,为建立非标准分析，只需要有一个模型就够了. 1972年美国的梵·奥士达 (Van Osdol)在《美国数学月刊》(Amer. Math. Monthly,79卷，355—363页)上发表了一篇文章，文中曾参照日本数学家高桥 (Takahashi) 的思想方法，具体地构造了一个较特殊的非标准数域模型,即所谓"超幂*R",或称为"高桥模型". 读者只需要知道集合论中的偏序集概念和"仓氏引理"(Zorn's Iemma)，就可以看懂*R 的构造方法,而不必求助于现代数理逻辑工具了.

为使推理更加直观简明，可借用等价过程量的表述形式来介绍超幂*R 的构造方法.

试考虑以 0,1,2,…为分段点的整点阶梯函数：

$$x(t)=x_n,\quad n\leqslant t<n+1,\ n\in \mathbf{Z}_0,$$

这里 $\mathbf{Z}_0$ 为非负整数集,$\{x_n\}$为通常的标准实数序列. 于是,由 $x(t)$作成的实数集

$$\tilde{x}=\{x(t)\mid 0\leqslant t<\infty\}$$

就叫作"过程量".特别,当 $x(t)=x_n\to 0$ (或∞)时，则相应的过程量 $\tilde{x}$ 叫作"趋零"(或"趋无穷")过程量. 又，如果 $x(t)=c$ (常数)，则相应的 $\tilde{x}$ 称为常过程量，可简记为 $\tilde{c}$.关于过程量之间的加、减、乘运算可定义为：

$$\tilde{x}\pm\tilde{y}=\{x(t)\pm y(t)\mid 0\leqslant t<\infty\},$$

$$\tilde{x}\cdot\tilde{y}=\{x(t)\cdot y(t)\mid 0\leqslant t<\infty\}.$$

只须规定 $\tilde{x}$ 的倒数概念,即可定义除法.如果 $x(t)$(其中 $0\leqslant t<\infty$)不恒为零,则可引入如下的整点阶梯函数 $y(t)$

(其中 $0\leqslant t<\infty$)：

$$y(t)=\begin{cases}0, & \text{当 } x(t)=0 \text{ 时;}\\ 1/x(t), & \text{当 } x(t)\neq 0 \text{ 时.}\end{cases}$$

由此引出的过程量 $\tilde{y}$ 即可定义为 $\tilde{x}$ 的倒数. 可以记为 $\tilde{y}=1/\tilde{x}$.

由如上所述，相信读者已能想象到各种“趋零”过程量当可用来规定无限小的概念. 但由于趋零速度有快有慢，这就必须把一切趋零过程量划分为许多“等价类”，使之表现许多不同的无限小. 换言之，各个无限小都要由相应的各个趋零过程量的等价类来代表. 这样，利用等价类概念，便可定义非标准实数——无限小. 当然，无限大便可通过非零无限小的倒数来定义，它们也有相应的过程量等价类. 一般说来，利用一般过程量的等价类，便可以界定广义实数(非标准实数与标准实数)的概念. 特别，常过程量的等价类即对应于标准实数. 因此，标准实数集 **R** 也就成为广义实数集 ***R** 的真子集了.

综上所述，所谓“高桥模型 ***R**” 就是由一切过程量等价类 $\langle x\rangle$ (即广义实数)作成的集合，即 $^*\mathbf{R}=\{\langle x\rangle\}$. 当然，我们必须要求 ***R** 成为一个有序数域. 因此问题归结为：应如何确立“等价性标准”，使之符合数学分析的需要？高桥与梵·奥士达的基本思想就是借用“超滤集” (Ultra-filter)概念，来规定过程量的等价性标准，以及它们是否具有某种性质的判别标准. 下面的论题属于纯粹抽象模式的构造分析，读者如感到集合论方面的准备知识不够，可以略而不读.

4. 超滤集的存在性

每一个数列 $\{x_n\}=\{x_0, x_1, x_2, \cdots\}$ 由映射 $\mathbf{Z}_0\mapsto\mathbf{R}$ 确定．如果知道幂集概念，可见$\{x_n\}$便是幂集 $\mathbf{R}^{\mathbf{Z}_0}$ 中的一成元．显然幂集 $\mathbf{R}^{\mathbf{Z}_0}$ 中的一切成元也就确定了一切过程量，这是因为有如下的一一对应关系：

$$\tilde{x} \leftrightarrow x(t) \leftrightarrow \{x_n\}, \quad \{x_n\}\in\mathbf{R}^{\mathbf{Z}_0}.$$

记 $x_n=f(n)$, $y_n=g(n)$, $z_n=h(n)$，等，则关于过程量 $\tilde{x}$、$\tilde{y}$、$\tilde{z}$ 的各种性质的表述，都可以转化为对 $f(n)$、$g(n)$、$h(n)$的性质表述．下面来引进滤集概念．

假设 $\mathbf{Z}_0$ 的某一子集类$\mathfrak{S}$具有下列三条性质：

1°空集 $\varnothing$ 不在$\mathfrak{S}$中．

2°如果 $S_1\in\mathfrak{S}$、$S_2\in\mathfrak{S}$，则它们的交集也在$\mathfrak{S}$中，即 $S_1\cap S_2\in\mathfrak{S}$.

3°如果 $S\in\mathfrak{S}$ 且 $S\subseteq T\subseteq\mathbf{Z}_0$，则 $T\in\mathfrak{S}$.

这样，$\mathfrak{S}$就称为 $\mathbf{Z}_0$ 上的一个"滤集"．

滤集的俗名叫"漏斗"．凡漏不出去的成元全体作成的集合就称"滤集"．例如上面的性质3°即表明，如果 S 漏不掉，则比它大的 T 也漏不掉．只有性质 2°颇为别致，它要求凡漏不去的诸成元之交集也漏不去．又由性质1°可知，滤集中不能有互不相交的成元．

［例］ 试考虑 $\mathbf{Z}_0$ 上的子集类：

$$\mathscr{F}=\{S\,|\,S\subseteq\mathbf{Z}_0,\ \mathbf{Z}_0\setminus S=\text{有限集}\}.$$

容易验明 $\mathscr{F}$ 是 $\mathbf{Z}_0$ 上的滤集，其中 $\mathbf{Z}_0\setminus S$ 表示差集(余集)．

为了对数列 $\{f(n)\}$、$\{g(n)\}$ 的属性能有一意性判

别标准，我们需要对 $\mathbf{Z}_0=\{n\}$ 的子集类作成的许多滤集中寻找出一个符合下列条件的尽可能大的"超滤集"：

如果滤集$\mathscr{U}$具有这样的性质：对 $\mathbf{Z}_0$ 的每一子集 S，或者是 $S\in\mathscr{U}$，或者是 $(\mathbf{Z}_0\setminus S)\in\mathscr{U}$，且两者中恰有一者成立，那末$\mathscr{U}$就叫作"超滤集".

超滤集就是这样的滤集，凡 $\mathbf{Z}_0$ 中的子集要么在该滤集中，要么其余集在滤集中，两者必居其一. 现在我们来证明一条存在性定理："至少存在一个 $\mathbf{Z}_0$上的超滤集，以$\mathscr{F}$为其子集."

证明 试取

$$\tilde{R}=\{\mathscr{F}'\mid\mathscr{F}'\supseteq\mathscr{F},\ \mathscr{F}'\text{ 为 }\mathbf{Z}_0\text{ 上的滤集}\},$$

即 $\tilde{R}$ 为 $\mathbf{Z}_0$ 上所有包含$\mathscr{F}$的滤集组成的类. 至少$\mathscr{F}$在其中，故 $\tilde{R}$ 为非空. 由于 $\tilde{R}$ 中诸滤集间的包含关系是一种偏序关系，并且 $\tilde{R}$ 中任一上升的链（即按序关系逐步增大的滤集序列）在 $\tilde{R}$ 中都有上界，这个上界就是链中各项(滤集)的并集. 因此，根据"仓氏引理"便得知 $\tilde{R}$ 中至少存在一个极大元$\mathscr{U}$(即它不是 $\tilde{R}$中任何元的子集)，当然它本身是个滤集. 往下我们要证明这个极大元$\mathscr{U}$正好就是一个超滤集.

采用反证法. 假设$\mathscr{U}$不是超滤集. 则由超滤集定义，可知至少有非空子集 $S\subseteq\mathbf{Z}_0$ 使 $S\notin\mathscr{U}$ 且 $(\mathbf{Z}_0\setminus S)\notin\mathscr{U}$. 这时，我们将能证明，必存在 $\mathbf{Z}_0$ 的一个子集 $T\in\mathscr{U}$，使得 $S\cap T=\varnothing$ (空集). 同理，将存在另一子集 $T'\in\mathscr{U}$ 使得 $(\mathbf{Z}_0\setminus S)\cap T'=\varnothing$. 既然 T 与 T' 分别和 S 及 $\mathbf{Z}\setminus S$不相交，则自然是 $T\subseteq\mathbf{Z}_0\setminus S$，$T'\subseteq S$. 由此又推得 $T\cap T'=\varnothing\in\mathscr{U}$.

注意$\mathscr{U}$本身是个滤集，而它含有$\varnothing$，这显然与滤集性质1°相矛盾．据此矛盾便推知原来假设不成立．故$\mathscr{U}$必为超滤集．

现在还须补证：$S\notin\mathscr{U}$，$(\mathbf{Z}_0\setminus S)\notin\mathscr{U}$ 条件下，必有$T\in\mathscr{U}$ 使得 $S\cap T=\varnothing$．再用反证法，假定 S 和$\mathscr{U}$的各成元 T 的交集都非空．试取

$$\mathscr{U}'=\{X\mid X\subseteq\mathbf{Z}_0,\ X\supseteq S\cap T,\text{对某个 } T\in\mathscr{U}\}.$$

可证 $\mathscr{U}'$ 也是一个滤集．因为 $S\cap T\neq\varnothing$，故滤集性质 1°和 3°是显然具备的．这里只须验明性质 2°．设 $X_1\in\mathscr{U}'$、$X_2\in\mathscr{U}'$，要求证明 $X_1\cap X_2\in\mathscr{U}'$．事实上，由 $\mathscr{U}'$ 的定义，知 $X_1\supseteq S\cap T_1$，$X_2\supseteq S\cap T_2$，T_1、$T_2\in\mathscr{U}$．注意 $\mathscr{U}$至少是个滤集，故知 $T_1\cap T_2\in\mathscr{U}$，从而

$$X_1\cap X_2\supseteq(S\cap T_1)\cap(S\cap T_2)=S\cap T_1\cap T_2.$$

这就证明了 $X_1\cap X_2\in\mathscr{U}'$．又因 $T\in\mathscr{U}\Rightarrow T\in\mathscr{U}'$，故 $\mathscr{U}\subseteq\mathscr{U}'$．最后，再注意 $S\supseteq S\cap T$，故 $S\in\mathscr{U}'$．但 $S\notin\mathscr{U}$，因此滤集 $\mathscr{U}'$ 真大于 $\mathscr{U}$．这和$\mathscr{U}$为极大元的性质矛盾，补证完成．至此，超滤集存在性定理便完全得证．

注意，以上所示只是一个纯粹存在性证明．由于 $\tilde{R}$中可以有许许多多的上升链，所以相应的极大元，也即超滤集，也可以有许许多多．至于每个超滤集的具体内容是什么，那是无法揭示的．当然它们都包含有 $\mathscr{F}$．事实上，为建立非标准数域$^*\mathbf{R}$ 的模型，只要有一个超滤集就够用了．

5．超滤集的作用

只要选定一个超滤集$\mathscr{U}$，就可用它来规定过程量的

种种性质，而不致于违反逻辑上的排中律．往后，我们将一律采用这样的定义：设有过程量 $\tilde{x}\leftrightarrow\{x_n\}$，$(n\in\mathbf{Z}_0)$．则称该过程量具有性质 P，当且仅当

$$\{n|x_n\text{ 具有性质 }P\}\in\mathscr{U}.$$

为说明这个定义的合理性，可以举一个简单的例子．比方说，P 代表一实数“是正的”这一性质．假定过程量 $\tilde{x}\leftrightarrow\{x_n\}$ 中的 x_n 对一切充分大的编号 n 恒为正（也即 $x_n>0$ 只要当 n 很大时）．于是

$$\{n|x_n>0\}\in\mathscr{F}\subseteq\mathscr{U}.$$

根据上述定义，就可以说“过程量 $\tilde{x}$ 是正的”．这当然是很合理的．

一般说来，如用 S、T 标记如下两个整数集

$$S=\{n|x_n\text{ 具有属性}P\},\quad T=\{n|x_n\text{ 不具有属性}P\},$$

则显然有 $S\cap T=\varnothing$，$S\cup T=\mathbf{Z}_0$，从而 $T=\mathbf{Z}_0\setminus S$．因此，根据超滤集 $\mathscr{U}$ 的本性，要么 $S\in\mathscr{U}$，要么 $T\in\mathscr{U}$，两者中恰有一者成立．这就是说，要么 $\tilde{x}$ 具有属性 P，要么它不具有属性 P，而两者中恰有一者为真．由此可见排中律总是适用的．这一点非常重要，否则在以$\mathscr{U}$为判别准则构成的数学模型（如$^*\mathbf{R}$）上就不能畅行无阻地进行逻辑推理了．

6．等价性概念与超幂$^*\mathbf{R}$

设 $\tilde{x}$ 与 $\tilde{y}$ 是两个过程量，且 $\tilde{x}\leftrightarrow\{x_n\}$，$\tilde{y}\leftrightarrow\{y_n\}$．如果 $\{n|x_n=y_n\}\in\mathscr{U}$，那末就说 $\tilde{x}$ 与 $\tilde{y}$互相等价，记作 $\tilde{x}=\tilde{y}$．

显然，如上规定的等价关系“$=$”具有反身性、对称性与传递性．这样，就可以按等价关系把 $\mathbf{R}^{\mathbf{Z}_0}$ 中的一切数列

进行分类．例如，同$\{x_n\}$等价的全体数列构成的类可以记作$\langle x\rangle$．由于$\{x_n\}\leftrightarrow\tilde{x}$，所以$\langle x\rangle$也就代表$\tilde{x}$的等价类．于是，所谓“超幂”$^*\mathbf{R}$就是由一切过程量的等价类$\langle x\rangle$作成的集合，也就是$^*\mathbf{R}=\{\langle x\rangle\}$．其中的$\langle x\rangle$就叫做“广义实数”．

注意，一个等价类$\langle x\rangle$（也即广义实数）可以由一个过程量$\tilde{x}$一意地确定．因此，任何一个等价类$\langle x\rangle$可以把它相应的一个$\tilde{x}$作为其代表．特别，对应于常过程量$\tilde{a}$，$\tilde{0}$，$\tilde{1}$的等价类可分别记作$\langle a\rangle$，$\langle 0\rangle$，$\langle 1\rangle$，等等．为简便计，也可用a，0，1等等来代表．这样，由所有常量c构成的标准实数集$\mathbf{R}=\{c\}$便可以看成是嵌入在$^*\mathbf{R}$内的一个真子集．

7．三分律及其验证

让我们来规定$^*\mathbf{R}$中成元（即广义实数）之间的大小关系：对于$^*\mathbf{R}$中的$\langle x\rangle$与$\langle y\rangle$，如果$\{n|x_n\leqslant y_n\}\in\mathscr{U}$，那末就规定$\langle x\rangle\leqslant\langle y\rangle$，并称$\langle x\rangle$**小于或等于**$\langle y\rangle$．

所谓$\langle x\rangle\neq\langle y\rangle$，意即指$\{n|x_n\neq y_n\}\in\mathscr{U}$．特别，如果$\langle x\rangle\leqslant\langle y\rangle$且$\langle x\rangle\neq\langle y\rangle$，则记作$\langle x\rangle<\langle y\rangle$．事实上这是因为（根据滤集性质2°）

$$\{n|x_n\leqslant y_n\}\cap\{n|x_n\neq y_n\}=\{n|x_n<y_n\}\in\mathscr{U}.$$

三分律就是这样一个命题：“对$^*\mathbf{R}$中的任何$\langle x\rangle$与$\langle y\rangle$，在下列三式中

$$\langle x\rangle=\langle y\rangle,\quad \langle x\rangle<\langle y\rangle,\quad \langle y\rangle<\langle x\rangle,$$

有且只有一式成立．”

证明　考虑如下三个集合（$\mathbf{Z}_0$的子集）：

$$E=\{n|x_n=y_n\},\quad L=\{n|x_n<y_n\},\quad G=\{n|y_n<x_n\}.$$

只有证明 E, L, G 中有且只有一个在$\mathscr{U}$中就可以了．首先，在标准数域 $\mathbf{R}$ 中三分律自然是成立的，即关系式 $x_n=y_n$，$x_n<y_n$，$y_n<x_n$ 三者中恰有一者成立．因此 $E\cup L\cup G=\mathbf{Z}_0$，且

$$E\cap L=E\cap G=L\cap G=\varnothing \text{（空集）}.$$

今证 E, L, G 中至少有一个是属于$\mathscr{U}$的．倘若 E 在$\mathscr{U}$中，则结论已真．否则，如果 $E\notin\mathscr{U}$，则因$\mathscr{U}$为超滤集，故 $\mathbf{Z}_0\setminus E=L\cup G\in\mathscr{U}$．此时，若 $L\in\mathscr{U}$，则结论已真．否则 $L\notin\mathscr{U}$，则 $\mathbf{Z}_0\setminus L=E\cup G\in\mathscr{U}$．于是，由滤集性质，知 $(L\cup G)\cap(E\cup G)=G\in\mathscr{U}$．这表明 E、L、G 三者中至少有一个属于$\mathscr{U}$．

又如果 E, L, G 中有某两者都属于$\mathscr{U}$，则按滤集性质，其交集（即空集 $\varnothing$）也将属于$\mathscr{U}$．这当然是不可能的，因此 E, L, G 中恰有一者在$\mathscr{U}$中，命题获证．

至此，我们已验明超幂$^*\mathbf{R}$ 是一个有序集．

8．作为有序域的$^*\mathbf{R}$

对$^*\mathbf{R}$ 中的元素，可规定加、减、乘运算如下：

$$\langle x\rangle\pm\langle y\rangle=\langle\tilde{x}\rangle\pm\langle\tilde{y}\rangle=\langle\tilde{x}\pm\tilde{y}\rangle=\langle x\pm y\rangle,$$

$$\langle x\rangle\cdot\langle y\rangle=\langle\tilde{x}\rangle\cdot\langle\tilde{y}\rangle=\langle\tilde{x}\cdot\tilde{y}\rangle=\langle x\cdot y\rangle.$$

容易看出，在这样定义之下的加法与乘法，是满足交换律、结合律和分配律的．又$\langle 0\rangle$，$\langle 1\rangle$和$\langle -x\rangle$分别是加法零元素、乘法单位元素和$\langle x\rangle$的关于加法的逆元素．至于关于乘法的逆元，也不难定义．今设$\langle x\rangle\neq\langle 0\rangle$，即$\{n|x_n\neq 0\}\in\mathscr{U}$．于是，由过程量 $\tilde{x}$ 的倒数 $\tilde{y}=1/\tilde{x}=\tilde{x}^{-1}$ 引出的

等价类$\langle\tilde{y}\rangle=\langle\tilde{x}^{-1}\rangle=\langle x^{-1}\rangle$，便可规定为$\langle x\rangle$的逆元，即$\langle x\rangle^{-1}=\langle y\rangle=\langle x^{-1}\rangle$．这显然是合理的，因为$\langle x\rangle\cdot\langle x\rangle^{-1}=\langle x\rangle\cdot\langle y\rangle=\langle x\cdot y\rangle=\langle\tilde{x}\cdot\tilde{x}^{-1}\rangle=\langle 1\rangle$．

综上所述，可知$^*\mathbf{R}$成为一个具有四则运算并且满足三分律的有序域．这就是按高桥思想构造出来的广义数域模型，而标准数域$\mathbf{R}=\{\langle c\rangle\}$作成$^*\mathbf{R}$的一个子域．往后，就称$^*\mathbf{R}$为非标准数域．

9．$^*\mathbf{R}$ 中的非标准实数

可以对$^*\mathbf{R}$中的元素引进绝对值的概念：

$$|\langle x\rangle|=\begin{cases}\langle x\rangle, & \text{当 } \langle x\rangle\geqslant\langle 0\rangle\text{时；}\\ -\langle x\rangle, & \text{当 } \langle x\rangle<\langle 0\rangle\text{时．}\end{cases}$$

为了写法简便，今后一律把$\langle x\rangle$，$\langle y\rangle$等记为x，y等，因而$|\langle x\rangle|$也可简记为$|x|$．

无限大与无限小便是$^*\mathbf{R}$中的“非标准实数”，其精确定义是：设$x\in{}^*\mathbf{R}$，如果对任意的正实数r，总成立着不等式关系$|x|<r$，则称x为**无限小**．又如果对任意的正实数r，总有$|x|>r$，则称x为**无限大**．假如存在一个固定正数r，使得$|x|\leqslant r$，则称x为**有限数**．凡无限小、无限大以及标准实数与无限小之和，都称为**非标准实数**．

读者不难观察到，由趋零过程量及趋无穷过程量作成的等价类(依照$\mathscr{U}$所定标准)，正好是符合上述定义的无限小与无限大．特别，$\langle 0\rangle$(或简记为0)也是一个无限小．

10．有限数的结构定理

凡是$^*\mathbf{R}$中的每一个有限数a，都可以唯一地表示成

这样的形式：

$$\alpha = \text{标准实数} + \text{无限小}.$$

这里所说的标准实数，是指子域 $\mathbf{R}=\{\langle c\rangle\}$中的常过程量的等价类.

证明 设 $\alpha\in{}^*\mathbf{R}$ 是有限数，我们要证明 α 可以表示成 $\alpha=r+\varepsilon$，其中 $r\in\mathbf{R}$ 而 ε 是个无限小.为此目的，我们来考虑下述集合

$$A=\{x\,|\,x<\alpha,\ x\in\mathbf{R}\}.$$

因为 α 是有限数，故有标准实数 $r_0\in\mathbf{R}$，使得 $\alpha<r_0$. 因此，A 是 $\mathbf{R}$ 中的有界集. 注意，由一一对应$\langle c\rangle\leftrightarrow c$ 可知 $R=\{\langle c\rangle\}$和经典分析中的通常实数域是同构的，所以 A 必有上确界 $d=\sup A$. 于是$|\alpha-d|$或为标准实数，或为非标准实数.

(1) 若$|\alpha-d|$是标准实数，即$|\alpha-d|=h\in\mathbf{R}$，则 $\alpha-d=\pm h$，也即 $\alpha=(d\pm h)+0$，0 也是无限小. 故定理成立.

(2) 若$|\alpha-d|$是非标准数，则可断言 $\alpha-d$ 为无限小. 假如不然，设$|\alpha-d|$不是无限小，则必存在某个正数 $h\in\mathbf{R}$ 使$|\alpha-d|>h$.于是可分两种情形：

(i) $\alpha-d>h$，这时 $\alpha>d+h$，故 $d+h\in A$.这说明 d 不是 A 的上确界，矛盾.

(ii) $d-\alpha>h$，这时 $d-h>\alpha$. 这也与 d 是 A 的上确界相矛盾.

因此 $\alpha-d$ 是无限小. 于是 $\alpha=d+(\alpha-d)$ 便是所要求的表达式（最后用反证法还可验明表达式的唯一

性). 故定理得证.

11. 有限数的标准部分

由结构定理，知$^*\mathbf{R}$中的有限数α均可表成标准实数与无限小之和，即

$$\alpha = {}^\circ\alpha + \varepsilon,\quad {}^\circ\alpha \in \mathbf{R},\ \varepsilon \text{ 为无限小}.$$

式中的标准实数${}^\circ\alpha$就叫作α的“标准部分”，有时也记作 st(α). (st 是 standard part 一词的缩写.)特别,α本身为标准实数时,则${}^\circ\alpha = \alpha$.

容易验明，对$^*\mathbf{R}$中的任意有限数α、β等，我们有下列简单等式

$$^\circ(\alpha+\beta) = {}^\circ\alpha + {}^\circ\beta,\qquad {}^\circ(\alpha-\beta) = {}^\circ\alpha - {}^\circ\beta,$$

$$^\circ(\alpha\cdot\beta) = {}^\circ\alpha\cdot{}^\circ\beta,\qquad {}^\circ\left(\frac{\alpha}{\beta}\right) = \frac{{}^\circ\alpha}{{}^\circ\beta}\ (\beta \text{ 不是无限小}).$$

特别,任何无限小ε的标准部分都为零,即${}^\circ\varepsilon = 0$. 对无限大,一般不定义标准部分.

如果我们取$x_n = n$, $n \in \mathbf{Z}_0$，则过程量$\tilde{x} \leftrightarrow \{n\}$的等价类可记作$\omega = \langle\tilde{x}\rangle$. 这是$^*\mathbf{R}\diagdown\mathbf{R}$中的一个无限大自然数. 显然，$n\omega$(这里$n \in \mathbf{Z}_+$)，$\omega^2 = \omega\cdot\omega$，$\omega^3 = \omega\cdot\omega\cdot\omega$，…等也都是无限大自然数,又可叫作“非标准自然数”. 全体无限大自然数作成集合$^*\mathbf{Z}$,它是$^*\mathbf{R}\diagdown\mathbf{R}$的子集. 通常把$\omega$叫作$^*\mathbf{Z}$中最小的无限大自然数. 自然，$\varepsilon = 1/\omega$便是一个无限小.

以上我们已经用了较多篇幅讨论了非标准数域$^*\mathbf{R}$(高桥模型)的构造. 这是极为有用的，因为有了这样一个包含着非标准实数的有序域，就可以在其上建立非标

准微积分了.

12. 牛顿"流数术"的重新解释

在前面(本书第 2 页)我们曾以函数 $f(t)=t^2$ 的求微商为例,说明了牛顿的流数术.现在我们采用非标准分析法,首先,必须把函数关系(或映射) $f:\mathbf{R}\mapsto\mathbf{R}$ 拓广到 $^*\mathbf{R}$ 上,即必须考虑扩张后的函数$^*f:{}^*\mathbf{R}\mapsto{}^*\mathbf{R}$. 具体写出来就是$^*f(^*t)={}^*t^2$, $(^*t\in{}^*\mathbf{R})$. 记 $^\circ(^*t)=t$,又令 ε 为无限小(可正可负). 于是 f 对 t 的微商即可由如下所示的标准部分得出:

$$^\circ\left(\frac{(^*t+\varepsilon)^2-{}^*t^2}{\varepsilon}\right)={}^\circ(2\cdot{}^*t+\varepsilon)=2t.$$

显然,牛顿的演算过程在形式上与此雷同.不过当时在他心目中并没有非标准数的概念,他更不会知道非标准数也可成为数域中的元素.

13. 扩张与对应置换

为了保证非标准分析方法能用以处理标准分析中的各种问题,就必须示明 $\mathbf{R}$ 上的一切数学分析"陈述"都可置换成$^*\mathbf{R}$ 上的相应"陈述",而且置换前后的陈述的真假性是一致的.这一原则称为"对应原理",它是由鲁宾逊首先提出的. 现今我们已经有了一个直观上比较显明的数域模型$^*\mathbf{R}$ (高桥模型), 因而在其上引入扩张与对应置换,其合理性也是相当明显的. 以下所述对应规则,均可作为定义看待.

首先,我们把 $\mathbf{R}$ 中的每一实数 a 都看成是常过程量的等价类$\langle a\rangle$,记作$^*a$,于是 $\mathbf{R}$ 中的 a 便扩展或置换成$^*\mathbf{R}$

中的 $^*a$，其对应置换关系可记作 $a\leftrightarrow{}^*a$.

"区间扩张"：给定 $\mathbf{R}$ 中的每一区间 (a, b)，对应地，可把它扩张成为 $^*\mathbf{R}$ 中的区间 $^*(a, b)=(^*a, ^*b)$. 其对应置换关系可记作 $(a, b)\leftrightarrow{}^*(a, b)$. 显然，$(a, b)$ 可以看作是嵌入在 $^*(a, b)$ 中的子集合. 同理，$\mathbf{R}$ 中的闭区间 $[a, b]$ 可以置换成 $^*\mathbf{R}$ 中的 $^*[a, b]=[^*a, ^*b]$.

"函数扩张"：给定标准函数 $f:\mathbf{R}\mapsto\mathbf{R}$，可以对应地扩张成为非标准函数 $^*f$: $^*\mathbf{R}\mapsto{}^*\mathbf{R}$，使得当 $x\in\mathbf{R}$ 时 $^*f(x)=f(x)$. 扩张的办法是这样的：令 $\tilde{x}$ 为由整点阶梯函数 $x(t)=x_n, n\leqslant t<n+1$, $n\in\mathbf{Z}_0$ 确定的过程量，即

$$\tilde{x}=\{x(t)\mid 0\leqslant t<\infty\}.$$

相应地可定义 $f(\tilde{x})=\{f(x(t))\mid 0\leqslant t<\infty\}$. 显然它还是一个阶梯式的过程量，因为 $f(x(t))$ 仍是 t 的整点阶梯函数. 今将过程量 $f(\tilde{x})$ 的等价类 $\langle f(\tilde{x})\rangle$ 记作 $^*f(\langle x\rangle)$，即

$$^*f(\langle x\rangle)=\langle f(\tilde{x})\rangle,$$

其中 $\langle x\rangle$ 是 $\tilde{x}$ 的等价类. 不难验明，这个定义实际上由 $\langle x\rangle$ 一意确定，而和类中的 $\tilde{x}$ 的选取无关.

特别地，如果让 $\langle x\rangle$ 遍取 $^*(a, b)$ 中的全体元素，则 $^*f(\langle x\rangle)$: $^*(a, b)\mapsto{}^*\mathbf{R}$. 可以证明，这 $^*f$ 就是标准函数 $f:(a, b)\mapsto\mathbf{R}$ 在 $^*(a, b)$ 上的一个扩张. 就是说，当 $x\in(a, b)$ 时 $^*f(x)=f(x)$. 为看出这一点，只须注意常过程量 $\langle c\rangle$ 可记作 c. 事实上，当常量 $x\in(a,b)$ 时，即有 $^*f(x)={}^*f(\langle x\rangle)=\langle f(\tilde{x})\rangle=\langle f(x)\rangle=f(x)$.

"关系扩张"：如上所示的函数扩张，可以拓广到一

般关系的扩张.即对于 $\mathbf{R}$ 上的每一个 n 元关系 $A(x_1,\cdots,x_n)$，总可扩张成 $^*\mathbf{R}$ 上的一个对应的关系 $^*A(x_1,\cdots,x_n)$，使得当各个 $x_i\in\mathbf{R}$ 时 $(i=1,\cdots,n)$，$^*A$ 与 A 相同，也即在 $\mathbf{R}$ 上有 $^*A(x_1,\cdots,x_n)=A(x_1,\cdots,x_n)$.

以后，当我们说到“在 $^*\mathbf{R}$ 上考虑”时，意即指有关的数学分析陈述(或问题)已经过扩张置换，变换成非标准分析的对应陈述(或问题). 又为了简便，凡扩张后的实数、点、区间、函数、关系等左上肩的 * 号都一律省略.(只要记住对应置换后它们已不是 $\mathbf{R}$ 上的对象.)

14. 微商概念

给定实变函数 $f:\mathbf{R}\mapsto\mathbf{R}$，并取定一点 $x_0\in\mathbf{R}$. 在 $^*\mathbf{R}$ 上考虑，令 ε 为无限小，如果下述标准部分为一有限实数：

$$^\circ\left(\frac{f(x_0+\varepsilon)-f(x_0)}{\varepsilon}\right)=f'(x_0),$$

则称 f 在 x_0 点可微，而 $f'(x_0)$ 便是 f 在 x_0 点的导数.

显然，在如上的导数概念中，已经包含着导数的非标准算法.从形式算法观点看，它是和牛顿的流数术一模一样的. 如果记 $\varepsilon=dx$，则 $f'(x_0)\varepsilon=f'(x_0)dx$ 就可作为微分 $df(x_0)$ 的定义.关于标准微分学中的许多公式和定理，当然都有相应的非标准分析的陈述形式，这里就不列举了.

15. 等和原理

初等积分学中的杜哈美(J. M. C. Duhamel)原理，在非标准分析中的对应物叫作“等和原理”. 它对处理无穷

小求和问题非常有用．

对*$\mathbf{R}$ 中的任意两个数 α 与 β，如果它们相差一个无限小，则称 α 与 β 彼此无限接近，记作 $\alpha\simeq\beta$．显然，凡无限小都无限接近于零．下面我们要用到无限大自然数 ω．

"等和原理"：假设在*$\mathbf{R}$ 上有两个正的无限小数列

$$\{\theta_1,\ \theta_2,\ \cdots,\ \theta_\omega\},\quad \{\theta'_1,\ \theta'_2,\ \cdots,\ \theta'_\omega\},$$

它们的各项之间有等价关系$\dfrac{\theta'_j}{\theta_j}\simeq 1$（其中 $j=1,\cdots,\omega$），且

$$\sum_{j=1}^{\omega}\theta_j=\theta_1+\theta_2+\cdots+\theta_\omega=l\ (\text{有限数}).$$

又设$\{r_1,\ r_2,\ \cdots,\ r_\omega\}$和$\{r'_1,\ r'_2,\ \cdots,r'_\omega\}$是两个有界数列，各项之间有关系 $r_j\simeq r'_j$（其中 $j=1,\cdots,\ \omega$）．那末成立如下的等和关系：

$$\sum_{j=1}^{\omega}r_j\theta_j\simeq\sum_{j=1}^{\omega}r'_j\theta'_j,\qquad \text{即}\qquad {}^{\circ}\Big(\sum_{j=1}^{\omega}r_j\theta_j\Big)={}^{\circ}\Big(\sum_{j=1}^{\omega}r'_j\theta'_j\Big).$$

这个原理的证明是不难的，这里从略．在应用此原理验证定积分的存在一意性时，往往只用它的特款 $\theta_j=\theta'_j$（正无限小）的情形．

16．积分概念及积分学基本定理

这里我们只以连续函数的积分为例，来表明非标准分析方法的特色．设给定一个连续函数 $f:[a,\ b]\mapsto\mathbf{R}$．今在*$\mathbf{R}$ 上考虑问题，任意取$[a,b]$的无限小划分方式

$$\Delta_\omega=\{a=x_0<x_1<\cdots<x_\omega=b\},$$

其中 $\rho_j=(x_j-x_{j-1})$都是无限小($j=1,\ 2,\ \cdots,\ \omega$)．再任意选定各点 $\xi_j\in[x_{j-1},\ x_j]$．于是，下列数值

$$^{\circ}\left(\sum_{j=1}^{\omega} f(\xi_j)(x_j-x_{j-1})\right)$$

是一意确定的，就是说，它和区间的无限小划分方式Δ_ω及选点 ξ_j 的方式都无关. 这个一意确定的数值，便定义为 f 在$[a, b]$上的**定积分**，即

$$\int_a^b f(x)dx = {}^{\circ}\left(\sum_{j=1}^{\omega} f(\xi_j)(x_j-x_{j-1})\right).$$

上述积分的存在一意性，可以用等和原理来证明，此处不细说了. 下面我们来证明牛顿与莱布尼茨的公式（积分学基本定理）

$$\int_a^b f(x)dx = F(b)-F(a),$$

这里 $F(x)$是 $f(x)$的原函数，即 $F'(x)=f(x)$，$(a<x<b)$.

证明 因为积分值是一意确定的，不妨就取一个等距的无限小划分方式 Δ_ω，其中 $x_j=a+(j/\omega)(b-a)$，$(j=1, 2, \cdots, \omega)$. 记 $x_j-x_{j-1}=(b-a)/\omega=dx$. 于是，根据导数的定义，并借助于等和原理，便得到

$$\begin{aligned}
F(b)-F(a) &= F(x_\omega)-F(x_0)\\
&= (F(x_1)-F(x_0))+(F(x_2)-F(x_1))\\
&\quad +\cdots+(F(x_\omega)-F(x_{\omega-1}))\\
&= \left(\frac{F(x_1)-F(x_0)}{x_1-x_0}\right)dx+\left(\frac{F(x_2)-F(x_1)}{x_2-x_1}\right)dx+\cdots\\
&\quad +\left(\frac{F(x_\omega)-F(x_{\omega-1})}{x_\omega-x_{\omega-1}}\right)dx\\
&\simeq F'(x_0)dx+F'(x_1)dx+\cdots+F'(x_{\omega-1})dx\\
&= \sum_{j=1}^{\omega} f(x_{j-1})\,dx
\end{aligned}$$

$$\simeq \sum_{j=1}^{\omega} f(x_j)\,dx$$

$$\simeq \int_a^b f(x)\,dx.$$

再注意,两个标准实数无限接近必然相等,这样就得到所要证明的公式.定理得证.

由上可以看出,正因为数域$^*\mathbf{R}$的建立过程早已经把极限概念融化进去了,所以只要在$^*\mathbf{R}$上考虑问题,就可以直接使非标准数(无限小与无限大)参加运算,从而也就无需再施加极限手续了.这显然是非标准分析方法的重要特色.事实上,微积分中有大量问题,按非标准分析法来处理都显得简捷明快.读者如欲了解更多题材,建议阅读鲁宾逊的原著.